

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN THỊ YẾN LINH

**NGHIÊN CỨU GIAO THỨC TRUYỀN GÓI TIN NGẮN CÓ ĐỘ TIN CẬY CAO
VÀ ĐỘ TRỄ THẤP TRONG TRUYỀN THÔNG CỘNG TÁC**

LUẬN ÁN TIẾN SĨ
NGÀNH: KỸ THUẬT ĐIỆN TỬ

TP. Hồ Chí Minh, tháng 06/2025

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN THỊ YẾN LINH

**NGHIÊN CỨU GIAO THỨC TRUYỀN GÓI TIN NGẮN CÓ ĐỘ TIN CẬY CAO
VÀ ĐỘ TRỄ THẤP TRONG TRUYỀN THÔNG CỘNG TÁC**

NGÀNH: KỸ THUẬT ĐIỆN TỬ

MÃ SỐ: 9520203

GVHD 1: PGS. TS. Phạm Ngọc Sơn

GVHD 2: GS. TS. Võ Nguyễn Quốc Bảo

TP. Hồ Chí Minh, tháng 06/2025

QUYẾT ĐỊNH GIAO ĐỀ TÀI LUẬN ÁN TIẾN SĨ

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

Số: 259/QĐ-ĐHSPKT

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

Tp. Hồ Chí Minh, ngày 02 tháng 02 năm 2021

QUYẾT ĐỊNH

Về việc giao đề tài luận án và người hướng dẫn NCS khóa 2020 - 2023

HIỆU TRƯỞNG TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH

Căn cứ Luật Giáo dục đại học ngày 18/6/2012 và Luật sửa đổi, bổ sung một số điều của Luật Giáo dục đại học ngày 19/11/2018;

Căn cứ Nghị định 99/2019/NĐ-CP ngày 30/12/2019 của Chính phủ Quy định chi tiết và hướng dẫn thi hành một số điều của Luật sửa đổi, bổ sung một số điều của Luật giáo dục đại học;

Căn cứ Quyết định số 937/QĐ-TTg ngày 30/6/2017 của Thủ tướng Chính phủ về việc phê duyệt đề án thí điểm đổi mới cơ chế hoạt động của Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh;

Căn cứ Nghị quyết số 11/NQ-HDT ngày 08/01/2021 của Hội đồng trường ban hành Quy chế tổ chức và hoạt động của Trường Đại học Sư phạm Kỹ thuật TP. HCM;

Căn cứ Thông tư số 08/2017/TT-BGDĐT ngày 04/4/2017 của Bộ Giáo dục và Đào tạo về việc Ban hành Quy chế tuyển sinh và đào tạo trình độ tiến sĩ;

Theo nhu cầu công tác và khả năng cán bộ;

Theo đề nghị của Trường khoa/Viện quản ngành và Trường phòng Đào tạo,

QUYẾT ĐỊNH:

Điều 1. Giao đề tài luận án tiến sĩ và người hướng dẫn cho:

Nghiên cứu sinh : *Nguyễn Thị Yến Linh*

Ngành : *Kỹ thuật điện tử*

Khoá: 2020 – 2023

Tên luận án : *Nghiên cứu giao thức truyền gói tin ngắn có độ tin cậy cao và độ trễ thấp trong truyền thông công tác*

Người HD thứ nhất (HD chính) : *TS. Phạm Ngọc Sơn*

Người HD thứ hai : *PGS.TS. Võ Nguyễn Quốc Bảo*

Thời gian thực hiện : *01/10/2020 đến 30/9/2023*

Điều 2. Giao cho Phòng Đào tạo quản lý, thực hiện theo đúng Quy chế đào tạo trình độ tiến sĩ của Bộ Giáo dục & Đào tạo và Nhà trường đã ban hành.

Điều 3. Trường các đơn vị, phòng Đào tạo, các Khoa/Viện quản ngành tiến sĩ và các Ông (Bà) có tên tại Điều 1 chịu trách nhiệm thi hành quyết định này.

Quyết định có hiệu lực kể từ ngày ký. 

Nơi nhận:

- BGH (để biết);
- Như điều 2, 3;
- Lưu: VT, ĐT (4b).



LÝ LỊCH CÁ NHÂN

I. Thông tin cá nhân:

Họ và tên: NGUYỄN THỊ YẾN LINH

Giới tính: Nữ

Ngày sinh: 04/08/1982

Nơi sinh: Tiền Giang

II. Quá trình đào tạo:

Từ 2000 - 2004: Sinh viên thuộc khoa Vật lý, Trường Đại học Khoa học Tự Nhiên, Tp.HCM.

Từ 2005 - 2008: Học viên cao học thuộc khoa Vật lý, Trường Đại học Khoa học Tự Nhiên, Tp.HCM.

Từ 2019 - hiện nay: Nghiên cứu sinh thuộc khoa Điện - Điện tử, Trường Đại học Sư phạm Kỹ Thuật Tp.HCM.

III. Quá trình công tác:

Từ năm 2009 đến hiện nay: Giảng viên khoa Cơ bản 2, Học Viện Công Nghệ Bưu Chính Viễn Thông Cơ sở tại Thành Phố Hồ Chí Minh.

TP. Hồ Chí Minh, ngày 16 tháng 06 năm 2025

Nguyễn Thị Yến Linh

LỜI CAM ĐOAN

Tôi cam đoan nội dung Luận án là công trình nghiên cứu của tôi.

Các số liệu và các kết quả nêu trong Luận án là trung thực và chưa được công bố trong bất kỳ công trình nghiên cứu nào khác.

TP. Hồ Chí Minh, ngày 16 tháng 06 năm 2025

Nguyễn Thị Yến Linh

LỜI CẢM TẠ

Đầu tiên, tôi xin chân thành cảm ơn GS. TS. Võ Nguyễn Quốc Bảo và PGS. TS. Phạm Ngọc Sơn đã hướng dẫn tận tình trong suốt quá trình học tập, nghiên cứu và hoàn thiện Luận án của tôi. Sự hỗ trợ quý báu và sự tận tâm của các Thầy đã giúp tôi vượt qua những thử thách, đồng thời nâng cao kiến thức và kỹ năng nghiên cứu của mình.

Tôi xin gửi lòng biết ơn đặc biệt đến các Thầy, Cô trong khoa Điện - Điện tử cũng như các Anh, Chị thuộc phòng Đào tạo sau Đại học của trường Đại học Sư phạm Kỹ thuật Tp. Hồ Chí Minh. Sự hỗ trợ và điều kiện tốt từ phía Trường đã giúp tôi có môi trường học tập và nghiên cứu thuận lợi, góp phần quan trọng vào quá trình hoàn thành Luận án của tôi.

Tôi cũng muốn bày tỏ lòng biết ơn chân thành đến Phó giám đốc Học Viện và các Thầy, Cô lãnh đạo Khoa Cơ bản 2 của Học Viện Công Nghệ Bưu Chính Viễn Thông Cơ sở tại Tp. Hồ Chí Minh. Sự giúp đỡ và tạo điều kiện thuận lợi từ phía Học Viện đã đóng góp không nhỏ vào quá trình học tập và nghiên cứu của tôi.

Cuối cùng, tôi muốn bày tỏ lòng biết ơn sâu sắc đến tất cả mọi người trong gia đình đã luôn đồng lòng động viên tinh thần, hỗ trợ vật chất và mang lại động lực lớn cho tôi suốt thời gian học tập và thực hiện Luận án. Sự ủng hộ từ gia đình là nguồn động viên quan trọng giúp tôi vượt qua khó khăn và hoàn thành chặng đường học tập của mình.

TP. Hồ Chí Minh, ngày 16 tháng 06 năm 2025

TÓM TẮT

Giao tiếp với độ tin cậy cao và độ trễ thấp (uRLLCs) là một yếu tố cốt lõi trong các mạng di động thế hệ 5G và 6G, đóng vai trò then chốt trong việc phát triển các ứng dụng Internet vạn vật (IoT) như tự động hóa công nghiệp, giao thông thông minh, y tế từ xa và thực tế ảo/tăng cường (VR/AR). Những ứng dụng này không chỉ góp phần nâng cao chất lượng cuộc sống mà còn thúc đẩy sự phát triển bền vững của nền kinh tế toàn cầu. Tuy nhiên, việc đáp ứng các yêu cầu nghiêm ngặt về độ tin cậy và độ trễ trong URLLCs đặt ra không ít thách thức trong thiết kế mạng không dây, đặc biệt khi các hệ thống này cần đảm bảo hiệu suất tối ưu trong môi trường thực tế. Công nghệ giao tiếp gói tin ngắn (SPC) được xem là một giải pháp quan trọng. Do đó, Luận án **“NGHIÊN CỨU GIAO THỨC TRUYỀN GÓI TIN NGẮN CÓ ĐỘ TIN CẬY CAO VÀ ĐỘ TRỄ THẤP TRONG TRUYỀN THÔNG CỘNG TÁC”** đề xuất ba mô hình mới cho truyền gói tin ngắn, nhằm cải thiện hiệu năng mạng, đáp ứng các yêu cầu khắt khe của uRLLCs, đồng thời tạo tiền đề cho sự phát triển mạng không dây thế hệ tiếp theo. Cụ thể:

Mô hình 1: Mô hình truyền gói tin ngắn trong mạng giải mã và chuyển tiếp (DF) hai chặng, trong đó, mô hình đã áp dụng các giao thức nhằm nâng cao độ phân tập không gian, bao gồm lựa chọn an-ten phát (TAS), lựa chọn nút chuyển tiếp một phần (PRS) và tỉ số kết hợp cực đại (MRC). Mô hình này so sánh hai giao thức chuyển tiếp SPC: giải mã và chuyển tiếp cố định (FDF) và giải mã và chuyển tiếp có chọn lựa (SDF). Kết quả cho thấy rằng hiệu năng của hệ thống FDF vượt trội hơn SDF ở mọi mức công suất phân bổ của nguồn phát. Đồng thời, mô hình cũng đề xuất bài toán tối ưu công suất phát và vị trí nút chuyển tiếp nhằm đáp ứng yêu cầu độ tin cậy và độ trễ cho hệ thống uRLLCs.

Mô hình 2: là mô hình truyền gói tin ngắn sử dụng mặt phản xạ thông minh (IRS) trong mạng chuyển tiếp. Mô hình này đề xuất phương pháp phân tích và đánh giá hiệu năng của hệ thống truyền gói tin ngắn qua IRS, so sánh với hệ thống truyền

thống sử dụng nút chuyển tiếp (DF và AF). Kết quả cho thấy, hệ thống SPC qua IRS đạt hiệu suất vượt trội so với các phương pháp truyền thống với cùng mức công suất phát, mở ra khả năng thay thế nút chuyển tiếp bằng IRS. Điều này không chỉ giúp cải thiện chất lượng dịch vụ (QoS) mà còn mở rộng phạm vi phủ sóng mạng với chi phí thấp hơn.

Mô hình 3: Mô hình truyền gói tin ngắn được hỗ trợ bởi mặt phản xạ thông minh IRS trong mạng cộng tác NOMA, trong đó, các giao thức được sử dụng nhằm tối ưu năng lượng, nâng cao độ tin cậy và giảm độ trễ cho hệ thống, thông qua việc định dạng sóng tại nguồn phát, mặt phản xạ thông minh và khử nhiễu liên tiếp (SIC) tại nguồn nhận. Để giảm thiểu thời gian tính toán và độ phức tạp, mô hình đề xuất phương pháp học sâu (DL) dựa trên mạng nơ-ron sâu (DNN) để dự đoán tỷ lệ lỗi bit (BLER) và công suất phân bổ nguồn phát. Các kết quả dự đoán này có ý nghĩa quan trọng trong việc phân bổ tài nguyên công suất một cách thông minh, hiệu quả, đáp ứng yêu cầu thời gian thực của hệ thống mạng thế hệ mới.

Những đóng góp của luận án này mang lại giá trị thiết thực trong việc phát triển mạng di động thế hệ mới, góp phần tạo nền tảng vững chắc cho các ứng dụng uRLLCs. Đồng thời, các phương pháp và mô hình mà luận án đề xuất có thể hỗ trợ hiệu quả cho việc thiết kế và triển khai mạng không dây trong tương lai.

ABSTRACT

Ultra-reliable and low-latency communications (uRLLCs) is a core element of 5G and 6G mobile networks, playing a pivotal role in the development of Internet of Things (IoT) applications such as industrial automation, smart transportation, remote healthcare, and virtual/augmented reality (VR/AR). These applications not only significantly enhance the quality of life but also drive the sustainable growth of the global economy. However, meeting the stringent requirements for reliability and latency in uRLLCs presents considerable challenges in the design of wireless networks, especially when these systems need to ensure optimal performance in real-world environments. Short packet communication (SPC) technology is considered a crucial solution.

Therefore, the dissertation titled "RESEARCH ON SHORT PACKET COMMUNICATION PROTOCOLS WITH ULTRA-RELIABLE AND LOW LATENCY IN COOPERATIVE COMMUNICATIONS" proposes three novel models for short packet transmission aimed at improving network performance, meeting the strict demands of uRLLCs, and providing a foundation for the development of next-generation wireless networks. Specifically:

Model 1: This model involves short-packet communications within a two-hop decode-and-forward (DF) network. The model incorporates protocols to enhance spatial diversity, including Transmit Antenna Selection (TAS), Partial Relay Selection (PRS), and Maximum Ratio Combining (MRC). Furthermore, the model concentrates on optimizing the transmit power and the position of the relay node to efficiently meet the service requirements of uRLLCs. The performance of the system is significantly improved in the optimized scenario compared to the non-optimized case.

Model 2: This model involves short-packet communications using Intelligent Reflecting Surfaces (IRS) in relay networks. It introduces a new approach to analyze, evaluate, and compare the system performance with that of the traditional

DF relay node short packet transmission models. This model achieves higher diversity order and performance improvements when compared with the DF relay model, despite both models using the same transmit power level.

Model 3: This model supports short-packet communications via Intelligent Reflecting Surfaces IRS within cooperative NOMA networks. It utilizes protocols aimed at optimizing energy, enhancing reliability, and reducing latency for the system, by shaping the waveforms at the source, through the intelligent reflecting surface, and with successive interference cancellation (SIC) at the receiver. Furthermore, this third model integrates deep learning methods based on deep neural networks (DNN) to predict performance and solve the power allocation problem, significantly improving accuracy while simultaneously reducing computation time and complexity compared to traditional methods.

The contributions of this dissertation provide practical value in the development of next-generation mobile networks, laying a solid foundation for uRLLC applications. Additionally, the methods and models proposed in this thesis can effectively support the design and deployment of future wireless networks.

DANH MỤC CÁC HÀM VÀ CÔNG THỨC TOÁN HỌC

STT	Ký hiệu	Ý nghĩa
1	$\arg \max(.)$	Hàm tìm một giá trị làm cho hàm số đạt giá trị lớn nhất.
2	$C(.)$	Dung lượng kênh Shannon với $C(x) = \log(1+x)$
3	$\exp(.)$	Lũy thừa cơ số e
4	$f_{\gamma}(.)$	Hàm mật độ xác suất (PDF)
5	$F_{\gamma}(.)$	Hàm phân phối tích lũy (CDF)
6	$\binom{.}{.}$	Hệ số nhị thức với $\binom{k}{n} = \frac{N!}{k!(N-k)!}$
7	$Q(.)$	Hàm Q-function với $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$
8	$V(.)$	Độ phân tán kênh truyền với $V(x) \triangleq \left(1 - \frac{1}{(1+x)^2}\right) (\log_2 e)^2$

9	$\mathbb{V}[\cdot]$	Toán tử phương sai
10	$\Gamma(\cdot)$	Hàm gamma với $\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt$
11	$\Gamma(\cdot, \cdot)$	Hàm gamma không hoàn chỉnh cận trên với $\Gamma(\alpha, x) = \int_x^{\infty} e^{-t} t^{\alpha-1} dt$
12	$\Upsilon(\cdot, \cdot)$	Hàm Gamma không hoàn chỉnh cận dưới với $\Upsilon(\alpha, x) = \int_0^x t^{\alpha-1} \exp(-t) dt$
13	$\mathbb{E}[\cdot]$	Toán tử kỳ vọng
14	$ \cdot ^2$	Toán tử độ lợi

MỤC LỤC

QUYẾT ĐỊNH GIAO ĐỀ TÀI LUẬN ÁN TIỀN SĨ.....	i
LÝ LỊCH CÁ NHÂN	ii
LỜI CAM ĐOAN	iii
LỜI CẢM ƠN.....	iv
TÓM TẮT	v
DANH MỤC CÁC HÀM VÀ CÔNG THỨC TOÁN HỌC.....	ix
MỤC LỤC.....	xi
DANH MỤC TỪ VIẾT TẮT.....	xv
DANH MỤC HÌNH VẼ.....	xviii
DANH MỤC BẢNG.....	xxi
MỞ ĐẦU	1
Bối cảnh nghiên cứu	1
Mục tiêu nghiên cứu	3
Nhiệm vụ nghiên cứu	4
Phạm vi nghiên cứu	4
Hướng tiếp cận và phương pháp nghiên cứu.....	4
Ý nghĩa khoa học và thực tiễn của đề tài nghiên cứu.....	5
Cấu trúc của Luận án.....	7
Chương 1. TỔNG QUAN	10
1.1 Giới thiệu.....	10

1.2 Các công bố quốc tế về truyền thông gói tin ngắn trong mạng cộng tác	10
1.3 Các công bố trong nước về truyền thông gói tin ngắn trong mạng cộng tác	16
1.4 Kết luận Chương 1	18
Chương 2. CƠ SỞ LÝ THUYẾT	19
2.1 Truyền thông tin gói ngắn	19
2.1.1 Cấu trúc gói tin ngắn.....	19
2.1.2 Đánh giá truyền thông tin gói ngắn.....	20
2.2 Truyền thông cộng tác	23
2.3 Mặt phản xạ thông minh.....	30
2.3.1 Cấu trúc IRS thụ động	30
2.3.2 Mô hình tín hiệu và kênh qua IRS	32
2.3.3 Ưu điểm của IRS	37
2.4 Kỹ thuật đa truy nhập phi trực giao miền công suất (PD-NOMA)	38
2.5 Kỹ thuật học sâu	43
Chương 3. TỐI ƯU HIỆU NĂNG HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG CHUYỀN TIẾP HAI CHẶNG	48
3.1 Giới thiệu	48
3.2 Mô hình nghiên cứu.....	49
3.3 Phân tích hiệu năng hệ thống	52
3.3.1 Tỷ lệ lỗi khối trung bình.....	52
3.3.2 Tỷ lệ lỗi khối dạng xấp xỉ	54
3.4 Tối ưu hiệu năng hệ thống.....	56

3.4.1 Tối ưu công suất nguồn phát.....	56
3.4.2 Tối ưu vị trí nút chuyển tiếp	59
3.5 Mô phỏng và thảo luận.....	63
3.6 Kết luận Chương 3	71
Chương 4. PHÂN TÍCH HIỆU NĂNG HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG CHUYỂN TIẾP SỬ DỤNG MẶT PHẢN XẠ	73
THÔNG MINH	73
4.1 Giới thiệu.....	73
4.2 Mô hình hệ thống	74
4.2.1 Hệ thống chuyển tiếp sử dụng mặt phản xạ thông minh	74
4.2.2 Hệ thống chuyển tiếp qua nút chuyển tiếp.....	80
4.3 Phân tích hiệu năng hệ thống	82
4.4 Mô phỏng và thảo luận.....	84
4.5 Kết luận Chương 4	87
Chương 5. THIẾT KẾ HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG ĐA TRUY NHẬP PHI TRỰC GIAO SỬ DỤNG MẶT PHẢN XẠ THÔNG MINH	89
5.1 Giới thiệu.....	89
5.2 Mô hình hệ thống.....	90
5.3 Phân tích hiệu năng hệ thống	93
5.3.1 Hàm CDF của SINR tại $\mathcal{U}_1$ và $\mathcal{U}_2$	93
5.3.2 Tỷ lệ lỗi khối.....	95

5.3.3 Tỷ lệ lỗi khối dạng xấp xỉ	99
5.3.4 Độ trễ và độ tin cậy	101
5.4 Mạng nơ ron sâu	101
5.4.1 Dự đoán BLER	102
5.4.2 Dự đoán công suất.....	107
5.5 Mô phỏng và thảo luận	108
5.5.1 Đánh giá hiệu năng.....	109
5.5.2 Dự đoán công suất	116
5.6 Kết luận Chương 5	118
Chương 6. KẾT LUẬN VÀ ĐỊNH HƯỚNG NGHIÊN CỨU TIẾP THEO	120
6.1 Giới thiệu	120
6.2 Kết luận của Luận án.....	120
6.3 Hướng phát triển của Luận án	122
PHỤ LỤC	124
TÀI LIỆU THAM KHẢO	134
DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA NCS	146

DANH MỤC TỪ VIẾT TẮT

Viết tắt	Tiếng anh	Tiếng việt
5G	Fifth Generation	Thế hệ thứ năm
6G	Sixth Generation	Thế hệ thứ sáu
AF	Amplify-and-Forward	Khuếch đại và chuyển tiếp
AI	Artificial Intelligence	Trí tuệ nhân tạo
AR	Augmented Reality	Thực tế tăng cường
AWGN	Additive White Gaussian noise	Nhiều trắng Gauss cộng tính
BLER	Block Error Rate	Tỉ lệ lỗi khối
CDF	Cumulative Distribution Function	Hàm phân phối tích lũy
CSI	Channel State Information	Thông tin trạng thái kênh
DF	Decode-and-Forward	Giải mã và chuyển tiếp
DNN	Deep Neuron Network	Mạng nơ ron sâu
DT	Direct Transmission	Truyền trực tiếp
eMBB	enhanced Mobile Broadband	Băng thông di động nâng cao
FDF	Full Duplex Relaying	Chuyển tiếp song công
FBL	Finite Blocklength	Chiều dài khối giới hạn
HDF	Half Duplex Relaying	Chuyển tiếp bán song công
H	Header	Phần đầu
i.i.d.	Independent & Identically Distributed	Phân bố độc lập và giống nhau

IFB	Infinite Blocklength	Chiều dài khối không giới hạn
IMT	International Mobile Telecommunications	Viễn thông Di động Quốc tế
IoT	Internet of Things	Kết nối vạn vật
ITU	International Telecommunication Union	Liên minh Viễn thông Quốc tế
IRS	Intelligent Reflect Surface	Mặt phản xạ thông minh
LMPs	Multilayer Neural Networks	Mạng nơ ron nhiều lớp
MAC	Media Access Control	Kiểm soát truy cập trung gian
MAPE	Mean Absolute Percent Error	Sai số tuyệt đối trung bình phần trăm
mMTC	massive Machine Type Communications	Giao tiếp loại máy lớn
ML	Machine Learning	Máy học
MIMO	Multiple Input Multiple Output	Đa đầu vào đa đầu ra
MRC	Maximal Ratio Combining	Kết hợp tỉ số cực đại
NOMA	Non-Orthogonal Multiple Access	Đa truy nhập phi trực giao
OP	Outage probability	Xác suất dừng
P	Preamble	Phần màu đầu
PA	Power Allocation	Phân bổ công suất
PHY	Physical layer	Lớp vật lý
PRS	Partial Relay Selection	Lựa chọn chuyển tiếp một phần
QoS	Quality of Service	Chất lượng dịch vụ

RMSE	Root Mean Square Error	Căn bậc hai sai số bình phương trung bình
SC	Selection Combining	Kết hợp lựa chọn
SL	Supervised Learning	Học có giám sát
RL	Relay Location	Vị trí nút chuyển tiếp
RV	Random variable	Biến ngẫu nhiên
SER	Symbol Error Rate	Tỉ lệ lỗi ký tự
SIC	Successive Interference Cancellation	Khử nhiễu tuần tự
SNR	Signal-to-Noise Ratio	Tỉ số tín hiệu trên nhiễu
SPC	Short Packet Communications	Truyền thông gói tin ngắn
STAR-IRS	Simultaneous Transmitting and Reflecting Intelligent Reflecting Rurfaces	Mặt phản xạ và truyền qua đồng thời thông minh
TAS	Transmit Antenna Selection	Lựa chọn an-ten phát
UL	Unsupervised Learning	Học không có giám sát
uRLLCs	Ultra-Reliable and Low-Latency Communications	Giao tiếp độ tin cậy cao và độ trễ thấp
V2X	Vehicle-to-Everything	Phương tiện kết nối mọi vật
VR	Virtual Reality	Thực tế ảo
WPIN	Wireless-Powered IoT Network	Mạng IoT hỗ trợ không dây

DANH MỤC HÌNH VẼ

Hình 2.1: Cấu trúc một gói tin ngắn.....	19
Hình 2.2: Sơ đồ khối minh họa cách tạo một gói tin.	21
Hình 2.3: Mô hình khối của một hệ thống truyền tin.....	21
Hình 2.4: Mô hình kênh của truyền thông cộng tác.....	24
Hình 2.5: Kỹ thuật chuyển tiếp DF.	25
Hình 2.6: Kỹ thuật chuyển tiếp AF.	26
Hình 2.7: Bộ kết hợp tuyến tính.....	28
Hình 2.8: Cấu trúc của mặt phản xạ thông minh IRS [27].	30
Hình 2.9: Điều chỉnh phần tử phản xạ dựa trên PIN diode.....	31
Hình 2.10: Mô hình tín hiệu qua IRS.....	36
Hình 2.11: Ước lượng kênh truyền trong mô hình IRS thụ động hoàn toàn (đường lên).	37
Hình 2.12: Mô hình truyền thông sử dụng PD-NOMA với K người dùng.	41
Hình 2.13: Sơ đồ Venn mô tả mối quan hệ giữa DL, ML và AI.	44
Hình 2.14: Cấu trúc của MLPs với 3 lớp ẩn.	44
Hình 2.15: Quá trình lan truyền tiến của DL.	46
Hình 2.16: Quá trình phản hồi lỗi ngược của DL.	47
Hình 3.1: Mô hình mạng DF hai chặng với các kỹ thuật TAS/PRS/MRC trong SPC. .	50
Hình 3.2: BLER toàn trình của SDF và FDF thay đổi theo SNR với $N = 2$ và.....	64
Hình 3.3: Sự ảnh hưởng e2e BLER vào số an-ten nguồn phát T với $N = 2$	65
Hình 3.4: e2e BLER theo số nút N và số an-ten M với ba trường hợp $\bar{\gamma} = 10, 15$ dB và 20 dB.	66
Hình 3.5: e2e BLER theo độ dài khối tin k với ba trường hợp $\bar{\gamma} = 10, 15$ dB và 20 dB	67

Hình 3.6: Sự phụ thuộc e2e BLER gần đúng và tiệm cận vào $\bar{\gamma}$ trong các trường hợp $k = 256, 512$ CUs và 1024 CUs .	68
Hình 3.7: Độ trễ và thông lượng toàn trình theo chiều dài khối k ở $\bar{\gamma} = 20$ dB với các trường hợp $\alpha = 0.5, 1.5$ và 2.5 .	69
Hình 3.8: e2e BLER theo SNR trung bình giữa trường hợp tối ưu và không tối ưu công suất.	70
Hình 3.9: e2e BLER theo SNR trung bình trong tình huống giữa tối ưu và không tối ưu vị trí.	71
Hình 4.1: Mô hình SPC được trợ giúp của IRS.	74
Hình 4.2. Hàm PDF của $\mathcal{A}$ với các giá trị L khác nhau.	80
Hình 4.3: Mô hình SPC thông qua chuyển tiếp FDF.	81
Hình 4.4: BLER giữa hai mô hình SPC qua IRS và qua $\mathcal{R}$ với $L = 2, L = 6$ và $L = 10$, tương ứng.	85
Hình 4.5: BLER theo số yếu tố phản xạ L trong hệ thống SPC qua IRS.	86
Hình 4.6: Sự ảnh hưởng BLER và độ trễ vào độ dài gói tin k trong hệ thống SPC qua IRS.	87
Hình 5.1: Mô hình hệ thống mạng NOMA-IRS.	91
Hình 5.2: Mô hình DNN cho dự đoán BLER tại $\mathcal{U}_i, i = \{1, 2\}$.	102
Hình 5.3: Giai đoạn huấn luyện và kiểm tra của mạng DNN.	104
Hình 5.4: Mô hình DNN cho dự đoán công suất tại $\mathcal{U}_i$	108
Hình 5.5: Sự tác động của SNR trung bình lên BLER tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với ba trường hợp hệ số $\alpha_1 = \alpha_2 = 0.5, \alpha_1 = 0.6, \alpha_2 = 0.4$ và $\alpha_1 = 0.8, \alpha_2 = 0.2$	110
Hình 5.6: Sự tác động của k lên BLER tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với $\alpha_1 = 0.6$ và $\alpha_2 = 0.4$	111
Hình 5.7: Sự tác động của L lên BLER tại $\mathcal{U}_2$ với ba trường hợp $\bar{\gamma} = 5$ dB, $\bar{\gamma} = 10$ dB và $\bar{\gamma} = 15$ dB .	112

Hình 5.8: Ảnh hưởng độ trễ và độ tin cậy lên k tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với trường hợp $\bar{\gamma} = 15$ dB và $\bar{\gamma} = 20$ dB .	113
Hình 5.9: RMSE và MAPE so với số lượng mẫu kiểm tra.	114
Hình 5.10: So sánh phần trăm độ chính xác của kết quả kiểm tra mô hình DNN tại $\mathcal{U}_1$ và $\mathcal{U}_2$.	115
Hình 5.11: BLER so với công suất phân bổ tại $\mathcal{U}_2$ ở các lớp khác nhau.	117
Hình 5.12: Ảnh hưởng RMSE và MAPE của DNN vào số lớp ẩn và số mẫu.	117

DANH MỤC BẢNG

Bảng 3.1: Thuật toán 1 tìm nghiệm gốc của (3.28).....	60
Bảng 3.2: Thuật toán 2 tìm nghiệm gốc của (3.38).....	62
Bảng 5.1: Tham số tính BLER.....	103
Bảng 5.2: Tham số huấn luyện DNN cho dự đoán BLER.....	109
Bảng 5.3: So sánh thời gian thực thi hệ thống trong ba phương pháp: phân tích lý thuyết, mô phỏng và mạng DNN tại $\mathcal{U}_1$ và $\mathcal{U}_2$	114
Bảng 5.4: Thời gian thực thi của mạng DNN, RMSE và MAPE theo số lớp ẩn khác nhau.....	118

MỞ ĐẦU

Bối cảnh nghiên cứu

Cho đến nay, các dịch vụ của Internet vạn vật (IoT), bao gồm công nghệ tự động hóa, hệ thống vận tải thông minh, Internet xúc giác, cũng như thực tế ảo (VR) và thực tế tăng cường (AR) đang phát triển nhanh chóng [1]. Công nghệ mạng không dây thế hệ thứ năm (5G) và các phiên bản tiên tiến hơn, bao gồm 5.5G và 6G, mở ra cánh cửa cho sự hội nhập sâu rộng của con người vào thế giới IoT và một xã hội siêu kết nối [2, 3]. Trong bối cảnh công nghệ mạng và truyền thông mới nổi, ba dịch vụ chính sẽ có những tiến bộ đột phá bao gồm giao tiếp máy với máy (mMTC), băng thông rộng di động cải tiến (eMBB), và giao tiếp độ tin cậy cao cùng độ trễ thấp (uRLLCs) [4, 5]. Theo tiêu chuẩn công nghệ của Viễn thông Di động Quốc tế IMT-2020/5G do liên minh Viễn thông Quốc tế (ITU-R) đề ra, công nghệ mạng 5G cam kết đạt được các chỉ tiêu về dung lượng truyền 10 – 100 Mbps/m², tốc độ truyền dữ liệu từ 10 – 20 Gbps, mật độ kết nối khoảng 10⁶ thiết bị/km² và độ trễ thấp từ 1 – 10 ms [6]. Hiện nay, việc phát triển dịch vụ uRLLCs, với đặc điểm là độ tin cậy cao và trễ thấp, đang trở thành ưu tiên chính trong ngành kỹ thuật và công nghiệp. Điều này tạo nền tảng vững chắc cho sự cải tiến của xã hội thông tin hiện đại [2, 7]. Truyền thông uRLLCs yêu cầu độ tin cậy cao đến 99.999% (tỷ lệ lỗi thấp hơn 10⁻⁵) và độ trễ ngắn (dưới 1 ms) nhằm đáp ứng các tiêu chuẩn cao về chất lượng dịch vụ (QoS) [8].

Gần đây, truyền thông gói tin ngắn (SPC) đã được nghiên cứu để cải thiện dịch vụ uRLLCs trong mạng 5G và các thế hệ mạng tương lai, thông qua việc giảm kích thước khung thông tin xuống còn 100 đơn vị sử dụng kênh (CUs) [9]. SPC đem lại sự đổi mới so với các hệ thống mạng trước đây sử dụng gói tin dài, với ưu điểm về độ trễ có thể được tính gần đúng trong một khe thời gian [10] và hiệu năng được đánh giá dựa trên tốc độ giải mã cực đại. Điều này là do các chỉ số hiệu năng của hệ thống truyền

thông tin truyền thống dựa trên lý thuyết thông tin cũ, áp dụng cho truyền tin qua gói dài. Do đó, việc nghiên cứu SPC hiện nay trở nên cấp thiết để đáp ứng đồng thời hai tiêu chuẩn quan trọng của hệ thống uRLLCs trong mạng 5G, bởi phần lớn các công trình nghiên cứu gần đây chỉ chú trọng giải quyết một trong hai yêu cầu, đó là độ tin cậy [11, 12]. Để hiện thực cho hệ thống uRLLCs, NCS đề xuất nghiên cứu các hướng tiềm năng như sau:

- Đầu tiên, việc tích hợp giao thức SPC vào mạng cộng tác đang được khám phá như một phương pháp tiềm năng để nâng cao hiệu năng của các hệ thống mạng thế hệ mới [12], [13]. Mạng cộng tác đã được kiểm chứng là phương pháp tăng cường hiệu quả công suất kênh truyền, độ tin cậy và mở rộng phạm vi phủ sóng [14, 15]. Ưu thế này đã được thể hiện rõ không chỉ trong các hệ thống mạng hiện tại mà còn trong việc phát triển các mạng thế hệ tương lai như 5.5G và 6G [16-18]. Tuy nhiên, các công trình nghiên cứu về SPC kết hợp kỹ thuật cộng tác thường chỉ nhấn mạnh vào việc cải thiện độ tin cậy so với hệ thống giao tiếp qua gói tin dài. Do đó, việc phát triển hệ thống mạng SPC hướng đến việc thỏa mãn cả tiêu chí về độ tin cậy và thời gian trễ, đặc biệt trong bối cảnh của các ứng dụng uRLLCs, là vấn đề nghiên cứu đáng được ưu tiên cao hiện nay.
- Thứ hai, trọng tâm nghiên cứu hiện nay là tăng cường sự phân tập của mạng. Do ràng buộc về độ trễ, việc triển khai phân tập theo thời gian trở nên khó khăn và không được tận dụng hiệu quả trong SPC. Như một giải pháp thay thế, việc sử dụng phân tập tần số và không gian, thông qua việc triển khai nhiều an-ten phát và nhận, là các giải pháp đáng chú ý. Các nghiên cứu mới đây đã chỉ ra rằng, việc triển khai các giao thức như tỉ số kết hợp cực đại (MRC) [12], chọn lựa nút chuyển tiếp một phần (PRS) [19], và hệ thống mạng đa đầu vào và đa đầu ra (MIMO) [20], đã phát huy hiệu quả trong việc cải thiện độ tin cậy vùng phủ và độ trễ của hệ thống SPC. Do đó, việc ưu tiên phát triển phân tập không gian trong thiết kế hệ thống SPC nhằm cải thiện đồng thời độ tin cậy và độ trễ là cần thiết hiện nay.

- Thứ ba, việc tối ưu phân bổ nguồn tài nguyên được xem là biện pháp cải thiện hiệu năng mà không cần tốn chi phí cho đa số hệ thống mạng truyền thông nhằm đáp ứng cả hai yêu cầu cơ bản: bảo đảm tính tin cậy và giảm độ trễ xuống mức thấp nhất. Then chốt của giải pháp này là tối ưu phân bổ công suất [21]. Điều này bao gồm việc tối ưu cách phân phối công suất truyền của các thiết bị trong mạng, với mục tiêu vừa đảm bảo hiệu suất truyền dữ liệu tốt nhất vừa tiết kiệm năng lượng. Đối với SPC, đặc biệt trong bối cảnh của hệ thống uRLLCs, việc điều chỉnh cả công suất và vị trí trở nên hết sức quan trọng, nhằm đảm bảo cả độ tin cậy và giảm thiểu độ trễ.

Trên các cơ sở đã nói trên, Luận án này định hướng đề xuất các giao thức bằng cách sử dụng các kỹ thuật tiên tiến ở lớp vật lý để giải quyết cả hai yếu tố độ tin cậy và độ trễ đối với SPC, nhằm hướng đến dịch vụ uRLLCs. Đề tài cụ thể của nghiên cứu là “Nghiên cứu giao thức truyền gói tin ngắn có độ tin cậy cao và độ trễ thấp trong truyền thông cộng tác”.

Luận án này đề cập đến hai phần chính: i) Tối ưu hiệu năng hệ thống SPC để phù hợp với yêu cầu của hệ thống uRLLCs trong mạng 5G; ii) Thiết kế các hệ thống SPC tích hợp các kỹ thuật tiên tiến hiện đại để nâng cao độ tin cậy và giảm thiểu độ trễ trong hệ thống mạng cộng tác.

Mục tiêu nghiên cứu

Trong Luận án này, nghiên cứu sinh (NCS) chú trọng vào hai mục tiêu chính được xác định rõ như sau:

- i. Mục tiêu thứ nhất tập trung vào việc xây dựng một mô hình SPC trong mạng truyền thông cộng tác, và thực hiện tối ưu hiệu năng của hệ thống nhằm đáp ứng các yêu cầu của hệ thống uRLLCs trong chuẩn mạng thế hệ mới.

- ii. Mục tiêu thứ hai nghiên cứu và áp dụng các phương pháp, kỹ thuật tiên tiến để cải thiện độ tin cậy và giảm thiểu độ trễ, cụ thể là áp dụng các kỹ thuật như mặt phản xạ thông minh, phương pháp đa truy nhập phi trực giao, và các kỹ thuật học sâu.

Nhiệm vụ nghiên cứu

Dựa trên hai mục tiêu chính của Luận án, NCS thực hiện những nhiệm vụ như đề cương đã được duyệt dưới đây:

- ✓ Phân tích và trình bày những khái niệm chính liên quan đến Luận án, bao gồm truyền tin gói ngắn, truyền thông cộng tác, mặt phản xạ thông minh (IRS), đa truy nhập phi trực giao (NOMA) và học sâu (DL).
- ✓ Để đạt được mục tiêu thứ nhất, hai nhiệm vụ được triển khai sau đây:
 - Nhiệm vụ 1: Xây dựng và đánh giá hiệu năng của mô hình SPC trong mạng chuyển tiếp hai chặng.
 - Nhiệm vụ 2: Tối ưu hiệu năng SPC trong mạng chuyển tiếp hai chặng để đáp ứng yêu cầu của hệ thống uRLLCs.
- ✓ Để thực hiện mục tiêu thứ hai, hai nhiệm vụ 3 và nhiệm vụ 4 được xác định là:
 - Nhiệm vụ 3: Phân tích hiệu năng của hệ thống SPC sử dụng IRS để nâng cao độ tin cậy và giảm độ trễ trong mạng cộng tác.
 - Nhiệm vụ 4: Ước lượng công suất phân bổ nguồn phát cho hệ thống SPC thông qua ứng dụng kỹ thuật DL, nhằm tuân thủ các tiêu chí của hệ thống uRLLCs.

Phạm vi nghiên cứu

Luận án này, NCS chủ yếu nghiên cứu SPC trong truyền thông cộng tác, tập trung vào mạng chuyển tiếp hai chặng và NOMA.

Hướng tiếp cận và phương pháp nghiên cứu

❖ Hướng tiếp cận

Căn cứ vào việc phân tích các công trình nghiên cứu trong và ngoài nước về SPC trong mạng thể hệ mới đến thời điểm viết Luận án, NCS đã tiến hành so sánh và đánh giá các kết quả cũng như các phương pháp từ những công trình này. Từ đó, NCS nhận định những vấn đề chưa được giải quyết và đưa ra các đề xuất cho hướng nghiên cứu tiếp theo trong của Luận án của mình.

❖ Phương pháp nghiên cứu

Trong Luận án này, NCS đã áp dụng các phương pháp nghiên cứu sau đây:

- ✓ Phương pháp phân tích và so sánh: NCS phân tích và so sánh tất cả nghiên cứu về SPC trong khuôn khổ của Luận án. Dựa trên đó, NCS đề xuất các mô hình SPC mới để đáp ứng mục tiêu nghiên cứu cụ thể của Luận án.
- ✓ Phương pháp phân tích thống kê bao gồm phân tích tín hiệu, toán giải tích, xác suất và toán tối ưu để tính toán các chỉ số đánh giá hiệu năng, cũng như giải quyết các vấn đề liên quan đến tối ưu hiệu năng cho các hệ thống được đề xuất.
- ✓ Phương pháp mô phỏng Monte Carlo được sử dụng để xác nhận và kiểm tra các kết quả phân tích lý thuyết.
- ✓ Phương pháp so sánh: NCS thực hiện so sánh hiệu năng của các mô hình hệ thống đề xuất so với các hệ thống hiện tại khác.

Ý nghĩa khoa học và thực tiễn của đề tài nghiên cứu

- Giá trị khoa học của Luận án này chủ yếu đóng góp cho khoa học các giải pháp, các kết quả tiên tiến về SPC, nhằm đáp ứng các tiêu chuẩn về độ tin cậy cao và độ trễ thấp trong các mạng không dây thể hệ mới như sau:
 - ✓ Luận án đưa ra các phương pháp để tăng cường độ tin cậy và giảm độ trễ trong SPC trong mạng chuyển tiếp hai chặng, bao gồm i) Tận dụng độ phân tập không gian thông qua việc áp dụng các kỹ thuật TAS, PRS và MRC; ii) Tối ưu hiệu năng hệ thống bằng cách xác định công suất phát và vị trí của nút chuyển tiếp, đảm bảo

tỉ lệ lỗi khối (BLER) tuân thủ các tiêu chuẩn uRLLCs mạng 5G, tức là $BLER \leq 10^{-5}$.

- ✓ Luận án này đưa ra giải pháp sử dụng IRS để thay thế cho nút chuyển tiếp trong hệ thống SPC. Với sự hỗ trợ của IRS, hệ thống SPC đạt được hiệu năng cải thiện hơn so với khi sử dụng nút chuyển tiếp DF và AF.
- ✓ Đồng thời, Luận án cũng lần đầu tiên đề xuất việc kết hợp kỹ thuật định dạng sóng với mặt IRS trong hệ thống mạng cộng tác NOMA, nhằm tăng cường độ tin cậy cho hệ thống SPC. Ngoài ra, để đáp ứng nhu cầu về độ trễ thấp và thời gian xử lý nhanh trong mạng tương lai, DL được áp dụng để giải quyết các thách thức liên quan đến dự đoán, bao gồm i) Các chỉ số đặc trưng cho hiệu năng như BLER, thông lượng và độ trễ; ii) Công suất phân bổ nguồn phát.

▪ **Ý nghĩa thực tiễn của Luận án**

- ✓ Mô hình đầu tiên, chuyên về SPC trong mạng chuyển tiếp hai chặng, được thiết lập đặc biệt cho mạng 5G, với các đặc điểm như công suất nguồn phát hạn chế, vùng phủ sóng hẹp, khoảng cách truyền ngắn và nhiều đồng kênh lớn. Hơn nữa, kết quả mới thu được từ mô hình này tạo ra một cơ sở vững chắc cho việc nghiên cứu và thiết kế tiếp theo của hệ thống mạng MIMO dành cho SPC.
- ✓ Các mô hình hỗ trợ SPC bằng mặt IRS đóng vai trò quan trọng trong cả mạng chuyển tiếp và mạng cộng tác NOMA. Những mô hình này hiệu quả trong mạng 5G và các mạng tương lai, đặc biệt trong các môi trường truyền khó khăn như: 1) Bị hạn chế về địa hình truyền sóng như đô thị và đồi núi; 2) Năng lượng nguồn phát bị hạn chế và vùng phủ sóng nhỏ.
- ✓ Kết quả từ Luận án đã được công bố trên các tạp chí quốc tế thuộc danh mục Web of Science (SCIE) như Journal of Communications and Networks (Q1) và Institution of Engineering and Technology Communications (Q2), cũng như tại Hội nghị quốc tế như 2023 RIVF International Conference on Computing and Communication Technologies (RIVF), 2024 International Conference on

Advanced Technologies for Communications (ATC). Đồng thời, kết quả của Luận án được xuất bản trên Tạp chí quốc gia nằm trong danh mục tạp chí của HDGS như Tạp chí Khoa học công nghệ Thông tin và Truyền thông (JSTICs). Những đóng góp này đã cung cấp nguồn tài liệu tham khảo chất lượng cho cộng đồng các nhà nghiên cứu, trung tâm và tổ chức nghiên cứu quốc tế trong lĩnh vực công nghệ mạng không dây mới.

Cấu trúc của Luận án

Cấu trúc của Luận án gồm phần mở đầu và sáu chương chi tiết được tổ chức như sau:

Phần mở đầu: Luận án đề cập đến bối cảnh nghiên cứu, các vấn đề chưa được giải quyết và nguyên nhân lựa chọn chủ đề nghiên cứu. Trong Phần này, mục tiêu nghiên cứu sẽ được đưa ra rõ ràng, dẫn đến việc xác lập các nhiệm vụ cốt lõi và phương pháp nghiên cứu cụ thể. Ngoài ra, đối tượng và phạm vi của nghiên cứu cũng được xác định rõ.

Chương 1: Tổng quan

Chương 1 này chủ yếu tập trung vào việc tổng hợp, phân tích và so sánh các nghiên cứu đã được công bố về SPC trong mạng cộng tác trong thời gian gần đây. Dựa trên những phân tích này, nghiên cứu sinh đã xác định những vấn đề còn tồn đọng, đồng thời xác định các nội dung có thể mở rộng và đưa ra các nhiệm vụ cụ thể cho Luận án.

Chương 2: Cơ sở lý thuyết

Chương 2 cung cấp một cái nhìn tổng quan về SPC và cơ chế truyền thông cộng tác. Chương 2 bao gồm việc trình bày về các chỉ số đánh giá hiệu năng của SPC cũng như giới thiệu các mô hình cơ bản trong truyền thông cộng tác. Chương này cũng phân tích ưu và nhược điểm của SPC và truyền thông cộng tác trong các hệ thống

uRLLCs. Ngoài ra, Chương 2 còn giới thiệu về các lý thuyết cơ bản liên quan đến các kỹ thuật cải tiến như IRS, NOMA, và DL.

Chương 3: Tối ưu hiệu năng hệ thống truyền thông tin gói ngắn trong mạng chuyển tiếp hai chặng

Chương 3 này phát triển mô hình giao tiếp SPC trong mạng DF hai chặng, được thiết kế để tích hợp nguồn phát và thu đa an ten, cũng như sự hỗ trợ của nhiều nút chuyển tiếp trung gian.

Điểm nổi bật của mô hình là việc áp dụng các kỹ thuật phân tập không gian như TAS, PRS và MRC. Hiệu năng hệ thống được phân tích qua BLER (cả dạng đóng và xấp xỉ), độ trễ và thông lượng toàn trình. Mô hình này đặc biệt nhấn mạnh việc tối ưu hiệu năng để đáp ứng các yêu cầu khắt khe của uRLLCs, bao gồm việc tối ưu công suất và vị trí của các nút chuyển tiếp, trong khi vẫn tuân thủ các giới hạn về tổng công suất và khoảng cách truyền cố định.

Nghiên cứu được trình bày trong Chương 3 đã góp mặt trong hai bài báo khoa học, cụ thể là bài số 2 và số 5, được liệt kê trong danh sách các công trình đã công bố của NCS.

Chương 4: Phân tích hiệu năng hệ thống truyền thông tin gói ngắn trong mạng chuyển tiếp sử dụng mặt phản xạ thông minh

Chương 4 đề cập đến việc đánh giá hiệu năng của hệ thống SPC trong môi trường mạng chuyển tiếp được hỗ trợ bởi công nghệ IRS. Phân tích hiệu năng tập trung vào việc đánh giá BLER, dưới cả hai dạng chính xác và xấp xỉ, theo các tham số hệ thống và kênh truyền như SNR, số lượng thành phần phản xạ, và độ dài khối tin.

Chương 4 này bao gồm kết quả từ công trình khoa học số 4 và số 3, thuộc danh sách công bố của NCS.

Chương 5: Thiết kế hệ thống truyền thông tin gói ngắn trong mạng đa truy nhập phi trực giao sử dụng mặt phản xạ thông minh

Chương 5 tập trung vào việc thiết kế hệ thống SPC trong mạng NOMA, được trợ giúp thông qua mặt IRS. Đặc biệt, nguồn đa an-ten sử dụng kỹ thuật tạo hướng sóng giúp điều hướng tín hiệu đến các người dùng mong muốn. Bên cạnh đó, Chương này cũng đánh giá hiệu năng của hệ thống dựa trên các yếu tố như BLER, độ trễ và độ tin cậy. Cuối cùng của Chương giới thiệu việc áp dụng DL để tối ưu việc phân bổ công suất trong hệ thống.

Chương 5 này cũng chính là nội dung được trình bày trong công trình số 1 của danh sách các công bố của NCS.

Chương 6: Kết luận và định hướng nghiên cứu tiếp theo

Trong Chương 6, Phần kết luận sẽ nêu bật các thành tựu và đóng góp chính của Luận án. Ngoài ra, Chương này cũng sẽ đề cập về các định hướng tiếp theo trong nghiên cứu dự kiến của NCS.

Chương 1. TỔNG QUAN

1.1 Giới thiệu

Chương 1 giới thiệu tổng quan những vấn đề có liên quan đến SPC trong mạng cộng tác. Những vấn đề này được đánh giá và phân tích dựa trên nghiên cứu từ các công trình quốc tế và trong nước liên quan trong những năm gần đây. Dựa trên cơ sở này, NCS đã xác định các khía cạnh còn tồn tại, tiềm năng phát triển và đồng thời đề xuất những nhiệm vụ cốt lõi trong Luận án. Bố cục của Chương 1 bao gồm phần khảo sát các công bố quốc tế và trong nước về SPC trong truyền thông cộng tác ở Mục 1.2 và 1.3. Phần kết luận của Chương 1 cũng được trình bày chi tiết ở Mục 1.4.

1.2 Các công bố quốc tế về truyền thông gói tin ngắn trong mạng cộng tác

Phần này, NCS liệt kê chi tiết các công bố quốc tế có liên hệ đến Luận án và tập trung vào từng nhiệm vụ cụ thể. Bên cạnh đó, NCS cũng trình bày rõ những vấn đề chưa giải quyết, giải pháp đưa ra cho những vấn đề này, cùng với các đối tượng và phương pháp nghiên cứu tương ứng cho từng nhiệm vụ như sau:

- ***Nhiệm vụ 1: Xây dựng mô hình và đánh giá hiệu năng hệ thống của mô hình truyền gói tin ngắn trong mạng chuyển tiếp hai chặng.***

Ban đầu, sự quan tâm của các nhà nghiên cứu đối với SPC trong truyền thông cộng tác được thể hiện qua các bài báo [11, 22]. Trong bài báo [11], các tác giả đã đề xuất một mô hình, trong đó nút nguồn truyền SPC với nút đích qua hai phương thức: trực tiếp (DT) và gián tiếp thông qua nút chuyển tiếp DF. Nút đích sử dụng kỹ thuật MRC để thu tín hiệu. Hơn nữa, hệ thống được mô tả dựa trên giả định rằng thông tin trạng thái kênh (CSI) chính xác tại cả nút nguồn và nút đích. Kết quả của [11] chứng minh rằng chuyển tiếp DF có tỉ số lỗi khối nhỏ hơn đáng kể so với truyền DT ở chế độ chiều dài khối tin ngắn (FBL). Mô hình giao tiếp SPC tương tự như trong bài báo [11], đã được triển khai trong [22]. Ở đây, phương pháp chuyển tiếp DF đã được chứng minh là nâng cao hiệu suất về thông lượng và hiệu

quả công suất so với phương thức truyền trực tiếp DT, đặc biệt trong các tình huống có CSI ở mức trung bình và hoàn hảo. Tiếp theo, nhóm tác giả trong [23] đã phát triển mô hình chuyển tiếp DF, bổ sung thêm các nút trung gian để nâng cao hiệu suất của kênh trong hệ thống truyền SPC, cả trong trường hợp CSI chính xác và khi CSI ở mức trung bình trong môi trường fading Rayleigh. Kết quả từ [23] cho thấy rằng, với CSI chính xác, thông lượng của hệ thống sử dụng mã hóa cho các khối tin ngắn FBL vượt trội so với hệ thống dùng mã hóa cho các khối tin dài (IBL).

Sau đó, nhóm tác giả trong [12] tiếp tục tập trung vào kịch bản chuyển tiếp DF hai chặng và mở rộng khai thác sự phân tập không gian vào hệ thống SPC thông qua các giao thức như kết hợp lựa chọn (SC) và MRC trên kênh fading Rayleigh bán tĩnh. Kết quả từ [12] đã chỉ ra rằng BLER giảm đáng kể theo SNR, bắt đầu với DT, tiếp theo là SC, và cuối cùng đạt đến giảm thiểu tối đa với MRC.

Tiếp theo, bài báo [24] đã phát triển mô hình từ nghiên cứu trong [12] bằng cách mở rộng kênh truyền fading Nakagami- m cho mô hình cộng tác DF. Kết quả từ [24] chứng tỏ rằng hiệu năng của hệ thống áp dụng giao thức MRC vượt trội hơn so với hệ thống dùng giao thức SC và DT, cả về công suất tiêu thụ lẫn đáp ứng yêu cầu độ trễ trong truyền thông M2M. Cuối cùng, nhóm tác giả trong [25] tiến hành mở rộng mô hình chuyển tiếp nhiều chặng trong SPC, sử dụng hai giao thức DF và AF, trong môi trường kênh fading Rayleigh bán tĩnh. Các kết quả phân tích và mô phỏng trong [25] đã cho thấy rằng độ trễ giải mã ảnh hưởng lớn đến thông lượng trong hệ thống chuyển tiếp SPC đa chặng. Khi độ trễ giải mã ở mức thấp, hệ thống sử dụng giao thức DF đạt thông lượng cao hơn so với AF. Đối với các hệ thống có số chặng chuyển tiếp vượt quá hai, thông lượng của DF nổi trội hơn AF ở mức độ trễ giải mã thấp và trung bình. Tuy nhiên, khi độ trễ giải mã tăng cao, thông lượng của cả hai DF và AF gần như tương đương nhau.

Tóm lại, các nghiên cứu về SPC trong mạng cộng tác chỉ mới bước đầu khai thác về độ tin cậy của hệ thống. Cụ thể, các phương pháp như DF và MRC được

tích hợp vào SPC để cải thiện hiệu năng cho hệ thống. Các bước tiếp cận bao gồm 1) Xây dựng mô hình SPC và áp dụng phân tích tín hiệu cùng lý thuyết xác suất để đánh giá BLER, thông lượng, và độ trễ; 2) Sử dụng mô phỏng để kiểm chứng các phân tích lý thuyết của mô hình. Tuy nhiên, trong bối cảnh các mạng không dây 5G và 6G, các tiêu chuẩn uRLLCs đặt ra yêu cầu cao về độ trễ và độ tin cậy, điều mà hầu hết các nghiên cứu trước đây chưa đạt được một cách toàn diện.

Trên cơ sở này, Luận án đề xuất nhiệm vụ 1 với mục tiêu đồng thời đảm bảo độ tin cậy và giảm độ trễ bằng cách khai thác sự phân tập về không gian của SPC. Cụ thể, hệ thống mới đề xuất tích hợp các kỹ thuật cải tiến về phân tập không gian như TAS, PRS và MRC trong mạng chuyển tiếp DF hai chặng.

✓ **Nhiệm vụ 2: Tối ưu hiệu năng của hệ thống truyền gói tin ngắn trong mạng chuyển tiếp cộng tác hai chặng theo yêu cầu của hệ thống uRLLCs.**

Tối ưu hiệu năng của hệ thống, hiện nay đang trở thành một thử thách trong nghiên cứu về SPC để đáp ứng yêu cầu của uRLLCs. Trong nghiên cứu gần đây, ở bài báo [24] nhóm tác giả đã chứng minh sự tương quan giữa độ tin cậy và độ trễ của SPC trong truyền thông cộng tác qua kênh fading Nakagami-m. Nhóm tác giả đề xuất thuật toán điểm nội (interior point algorithm) để tối ưu hàm mục tiêu không tuyến tính của xác suất dừng (OP) để đạt được giá trị tối ưu cho công suất. Kết quả của nghiên cứu chứng tỏ khi hệ thống hoạt động với công suất tối ưu thì độ trễ giảm đáng kể.

Mặc dù, vấn đề tối ưu hệ thống mạng không là điều mới lạ trong các hệ thống truyền thông truyền gói tin dài, nhưng đối với SPC, vấn đề này ngày càng quan trọng để bảo đảm thỏa mãn hoàn toàn các tiêu chuẩn về độ tin cậy và về thời gian trễ cho các hệ thống mạng thế hệ mới. Hiện tại, vấn đề tối ưu trong SPC vẫn đối mặt với những khó khăn đáng kể, chủ yếu bởi vì việc nghiên cứu SPC đòi hỏi phải thực hiện các phép tích phân phức tạp và công thức tính BLER thường không tuân theo hình dạng lồi lõm. Điều này dẫn đến thực tế là gần như không tồn tại một phương pháp chung nào để giải quyết các vấn đề tối ưu với hàm không lồi.

Trên cơ sở này, Nhiệm vụ 2 tập trung vào việc cải thiện hiệu năng cho hệ thống mạng DF hai chặng ở kênh truyền fading Rayleigh. Điều này bao gồm việc tối ưu vị trí của nút chuyển tiếp và công suất cho nút nguồn, trong khi tuân thủ các hạn chế như khoảng cách hoặc tổng công suất cho trước.

Nhiệm vụ 2 này mang lại hai kết quả chính như: i) Về lý thuyết, xác định được khoảng cách tối ưu để đặt nút chuyển tiếp và công suất phát tối ưu; và ii) Về mô phỏng, kiểm tra lại tính đúng đắn của các kết quả lý thuyết.

✓ ***Nhiệm vụ 3: Phân tích hiệu năng hệ thống truyền thông tin gói ngắn qua mặt phản xạ thông minh nhằm nâng cao độ tin cậy và giảm độ trễ cho hệ thống mạng cộng tác.***

Nhiệm vụ 3 được xây dựng nhằm thực hiện mục tiêu số 2 trong Luận án, thông qua việc khảo sát sâu rộng những công trình bên dưới như sau:

- *Các nghiên cứu liên quan*

Gần đây, công nghệ IRS đã được giới thiệu như một giải pháp mới, hứa hẹn mang lại môi trường thông minh cho các mạng di động thế hệ mới 5G và 6G [26, 27]. Điểm nổi bật của môi trường vô tuyến thông minh so với các mạng không dây truyền thống là khả năng điều chỉnh không gian thông qua phần mềm, mang lại sự linh hoạt và kiểm soát cao [28].

Hiện nay, các nghiên cứu đã phát hiện ra rằng IRS mang lại nhiều lợi ích đáng kể trong việc truyền và điều chỉnh thông tin. Thứ nhất, IRS chi phí thấp và có thể lắp đặt trên nhiều loại bề mặt khác nhau, bao gồm mặt ngoài của tòa nhà, cột trên cao, panô quảng cáo [29]. Thứ hai, IRS nâng cao hiệu quả sử dụng băng tần bằng cách tái cấu hình môi trường truyền không dây. Cụ thể, IRS giúp tăng cường sự phủ sóng và cải thiện hiệu suất truyền dẫn tín hiệu trong các khu vực có tín hiệu yếu [30]. Thứ ba, IRS giúp giảm lượng năng lượng tiêu thụ và đem lại lợi ích cho môi trường, nhờ khả năng điều chỉnh pha của các thành phần phản xạ để hình thành tín hiệu mà không cần dùng đến bộ khuếch đại công suất [31, 32]. Cuối cùng, việc tích hợp IRS vào mạng không dây tăng cường tính tương thích với tiêu chuẩn và phần cứng của hệ thống mạng hiện hành [33].

Trên cơ sở các lợi ích của IRS trong hệ thống truyền gói tin dài ở trên, gần đây IRS cũng được chứng minh là giải pháp hứa hẹn đáp ứng yêu cầu uRLLCs trong các hệ thống tự động hóa nhà máy. IRS tăng cường hiệu quả truyền thông trong các hệ thống giao tiếp máy với máy như trong bài báo [34]. Nhóm tác giả của bài báo [34] tập trung vào việc phân tích tốc độ dữ liệu trung bình và tỉ lệ lỗi giải mã trung bình qua các kênh như fading Nakagami- m và fading Rician. Một cách cụ thể hơn, trong bài báo [35] nhóm tác giả tiếp tục nghiên cứu IRS hỗ trợ trong SPC trong hệ thống mạng NOMA. Ở đây, nhóm tác giả dựa vào BLER trung bình và thông lượng để đo lường hiệu năng của các người dùng trong mạng NOMA.

Gần đây, sự quan tâm nghiên cứu IRS hỗ trợ các hệ thống SPC đã tăng lên với các kết quả đầu tiên chứng minh hiệu quả nổi bật của IRS. Tuy nhiên, hiện nay nghiên cứu IRS hỗ trợ SPC trong hệ thống mạng cộng tác vẫn còn hạn chế. Trên cơ sở này, Nhiệm vụ 3 nghiên cứu hiệu năng của hệ thống chuyển tiếp được trợ giúp bởi IRS trong SPC. Các phương pháp đề xuất bao gồm phân tích tính hiệu, xác suất thống kê để tính BLER và độ lợi phân tập hệ thống. Tiếp theo, phương pháp mô phỏng Monte Carlo cũng được áp dụng xác minh lại các kết quả của phân tích lý thuyết. Cuối cùng, nghiên cứu thực hiện đánh giá và so sánh hiệu năng giữa mô hình SPC thông qua IRS với mô hình SPC qua nút chuyển tiếp DF.

Mục tiêu của nhiệm vụ thứ 3 là xác nhận sự cải thiện hiệu năng của IRS trong việc truyền SPC so với chuyển tiếp DF thông thường.

✓ *Nhiệm vụ 4: Ước lượng công suất phân bổ nguồn phát cho hệ thống truyền gói tin ngắn trong mạng truyền thông cộng tác bằng kỹ thuật học sâu nhằm đáp ứng yêu cầu uRLLCs.*

Dựa vào các nghiên cứu sâu rộng SPC về các công nghệ mới chuẩn bị cho mạng 6G, Nhiệm vụ 4 được đưa ra nhằm thực hiện mục tiêu thứ 2 trong Luận án như sau:

- *Các nghiên cứu liên quan*

Gần đây, 6G được dự kiến sẽ tạo ra những đột phá quan trọng, không chỉ giới hạn ở kết nối di động mà còn mở rộng sang việc tích hợp trí tuệ nhân tạo (AI) thông qua cải tiến các giao thức và kiến trúc mạng. Thực tế, AI đã giúp giải quyết hầu hết các tình huống khó khăn như điều phối lưu lượng mạng, điều chỉnh kết nối, an ninh và bảo mật thông tin [36]. Không giới hạn ở việc nghiên cứu sâu rộng trong ngành khoa học máy tính, hiện tại AI đang nhận được sự chú ý từ cộng đồng nghiên cứu thông tin và mạng vô tuyến. Đặc biệt, DL một nhánh quan trọng của máy học (ML) thuộc lĩnh vực mạng nơ ron sâu (DNN), đang được chú trọng nhằm hướng đến đáp ứng yêu cầu cho uRLLCs. Cụ thể, một số công trình ban đầu đã chứng tỏ ưu thế của DL như bài báo [37]. Nhóm tác giả đã đưa ra một thuật toán học củng cố để tối ưu lịch truyền trong uRLLCs, nhằm cải tiến QoS. Hay ở bài báo [38, 39] đã chứng minh rằng DL không có giám sát đáp ứng được QoS về độ trễ và độ tin cậy cho người dùng trong mạng uRLLCs. Gần đây, DL được chú ý nghiên cứu trong SPC như trong bài báo [40]. Cụ thể ở bài báo [40], nhóm tác giả đã đưa ra DL để dự đoán sự phân bố công suất nguồn và lập lịch truy cập uRLLCs cho hệ thống SPC. Kết quả bài báo cho thấy DL mang lại kết quả không chỉ cung cấp dự đoán chính xác mà đáp ứng yêu cầu về độ trễ thấp, phù hợp với tiêu chuẩn của hệ thống 5G. DL được vận dụng giải quyết bài toán dự đoán nguồn công suất phân bố cho hệ thống và lập lịch truy cập hệ thống.

Hiện nay, nghiên cứu DL hỗ trợ trong SPC đang là vấn đề mới, cụ thể số lượng công bố khoa học về chủ đề này còn khá hạn chế. Do đó, trong Nhiệm vụ 4 này DL kết hợp SPC nhằm giải quyết những tồn tại khi sử dụng các phương pháp truyền thống như 1) Áp dụng cho mọi bài toán và bất kể độ phức tạp; 2) Thời gian mô phỏng nhanh, vượt trội so với phương pháp Monte Carlo, và phù hợp với yêu cầu thời gian thực thi ngắn trong các hệ thống mạng tương lai; 3) Thực hiện cho mọi bài toán, từ bài toán thuận như dự đoán hiệu năng hệ thống đến bài toán nghịch là tối ưu, tức là dự đoán công suất nguồn phát dựa trên BLER cho trước, thỏa mãn yêu cầu uRLLCs.

1.3 Các công bố trong nước về truyền thông gói tin ngắn trong mạng cộng tác

Hiện nay, mục tiêu cải tiến hiệu năng mạng và tiến đến tăng cường QoS để tuân thủ các yêu cầu của uRLLCs đang trở thành mục tiêu chính của gần như tất cả hệ thống mạng 5G và 6G sắp tới. Nghiên cứu SPC trong mạng cộng tác nhằm hiện thực uRLLCs được coi là một phương án triển vọng đang được những nhà khoa học chú trọng.

❖ *Giao tiếp gói tin ngắn trong mạng chuyển tiếp cộng tác*

Đầu tiên, nhóm tác giả trong bài báo [19] nghiên cứu về SPC trong hệ thống mạng DF hai chặng ở kênh fading Rayleigh. Đặc biệt, phương pháp PRS được áp dụng để chọn một nút chuyển tiếp tốt nhất từ tập đa nút chuyển tiếp trung gian. Bài báo cho thấy rằng BLER giảm nhanh khi tăng các yếu tố như tỉ lệ tín hiệu trên nhiễu (SNR), số lượng nút chuyển tiếp, và chiều dài của khối tin. Tiếp theo, nhóm tác giả trong các bài báo [41, 42] đã phát triển mô hình SPC để áp dụng cho chuyển tiếp DF qua nhiều chặng trên kênh fading Rayleigh. Đặc biệt, bài báo [41] khẳng định lại rằng hiệu năng được cải tiến khi tăng công suất phát và chiều dài khối tin. Quan trọng hơn hết đối với mạng DF cho SPC là hiệu năng hệ thống giảm khi số chặng tăng, và đạt tối ưu với hai chặng. Tuy nhiên, trong bài báo [42] nhóm tác giả tận dụng truyền thông cộng tác tăng cường tại mỗi chặng để nâng cao hiệu năng trong mạng đa chặng SPC. Các kết quả của [42] cho thấy rằng, bằng cách áp dụng sự tăng cường cộng tác ở mỗi chặng, hiệu năng của hệ thống SPC đã được cải thiện đáng kể so với mô hình truyền thông cộng tác truyền thống.

Nói chung, các nhóm nghiên cứu trong nước đã bắt đầu quan tâm đến SPC trong mạng cộng tác. Hầu hết nghiên cứu tập trung phân tích ảnh hưởng của BLER vào SNR, chiều dài khối tin, số chặng, và số nút chuyển tiếp. Hiện nay, thách thức vẫn còn tồn tại là làm sao để các hệ thống SPC trong mạng cộng tác tuân thủ các tiêu chuẩn của uRLLCs, điều này vẫn là một lĩnh vực mở cần được nghiên cứu và giải quyết thêm.

❖ *Mặt phản xạ thông minh hỗ trợ truyền gói tin ngắn trong truyền thông cộng tác*

Gần đây, các nhà khoa học trong nước đã bắt đầu quan tâm đến IRS nhằm tăng cường độ tin cậy trong giao tiếp SPC như sau: 1) IRS hỗ trợ trong các hệ thống vô tuyến nhận thức dạng nền qua bài báo [43]; 2) IRS hỗ trợ cho hệ thống thiết bị bay không người lái qua bài báo [44].

Tuy nhiên, theo tìm hiểu của NCS, nghiên cứu ứng dụng IRS để hỗ trợ hệ thống SPC trong truyền thông cộng tác hiện nay vẫn chưa được đề ý đến. Do đó, đề tài khám phá IRS với mục tiêu cải thiện đồng thời độ tin cậy và giảm độ trễ cho hệ thống SPC hiện tại đang là chủ đề mở đối với cộng đồng khoa học trong nước.

❖ *Phương pháp học sâu hỗ trợ truyền thông tin gói ngắn trong truyền thông cộng tác*

Gần đây, nhóm nghiên cứu trong nước bắt đầu đã chú ý đến việc tích hợp DL vào SPC, nhằm đảm bảo về độ trễ trong các thiết bị IoT của các mạng thế hệ mới, được thể hiện qua các bài báo [45, 46]. Cụ thể, bài báo [45] nghiên cứu mạng IoT-NOMA nhận thức SPC để cải thiện hiệu quả vùng phủ, đồng thời giảm độ trễ dưới điều kiện CSI không hoàn hảo và khử nhiễu tuần tự (SIC). Nhóm tác giả đã thực hiện: 1) Phân tích tỉ lệ lỗi khối BLER và thông lượng của các người dùng trong mạng; 2) Đề xuất mô hình DL để dự đoán đồng thời BLER và thông lượng của các người dùng. Kết quả cũng chứng tỏ rằng hiệu quả dự đoán từ DL có tỉ lệ lỗi thấp và thời gian dự đoán ngắn hơn các mô hình học khác. Bài báo [46], nhóm tác giả cũng đề ý đến mạng IoT thu thập năng lượng, trong đó các thiết bị IoT truyền gửi gói tin ngắn SPC đến nhiều nút đích bằng cách thu năng lượng từ nhiều bộ phát qua kênh fading Rayleigh. Đồng thời, nhóm tác giả cũng phân tích BLER và thông lượng của hệ thống. Ngoài ra, nhóm tác giả thiết kế mạng DNN để dự đoán thông lượng hệ thống dưới các cài đặt kênh khác nhau. Kết quả bài báo chứng minh mô hình DNN có khả năng dự đoán thông lượng và thời gian thực thi một cách chính xác, vượt trội so với các phương pháp phân tích số và mô phỏng.

Nhìn chung, nhóm tác giả nghiên cứu tích hợp DL vào các hệ thống SPC đã cho thấy tiềm năng ban đầu trong việc giải quyết vấn đề giảm thiểu độ trễ của hệ thống mạng theo chuẩn uRLLCs cho các hệ thống IoT. Nhóm nghiên cứu tập trung vào DL để dự đoán các chỉ số hiệu năng như BLER và thông lượng mạng. Tuy nhiên, các nhóm chỉ giải quyết một chiều về thời gian trễ còn độ tin cậy của hệ thống thì vẫn chưa tập trung khai thác. Cụ thể, vấn đề tối ưu hóa hiệu năng hệ thống dựa vào DL để cải tiến độ tin cậy vẫn chưa được giải quyết được.

1.4 Kết luận Chương 1

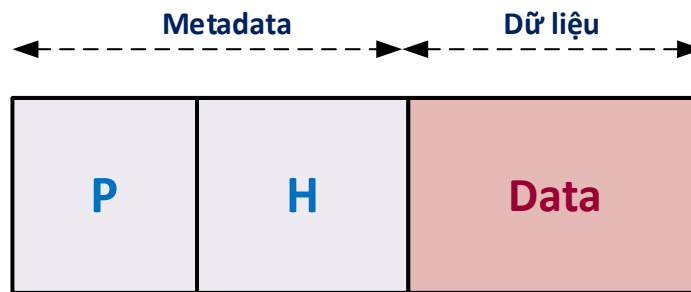
Trong Chương 1, NCS đã phân tích và so sánh chi tiết các nghiên cứu liên quan SPC, cả trong nước và quốc tế liên quan đến các nhiệm vụ cụ thể trong phạm vi Luận án. Qua đánh giá sâu rộng của các công trình này, NCS đã đề xuất bốn nhiệm vụ cơ bản để thực hiện hai mục tiêu chính của Luận án. Cụ thể, Nhiệm vụ 1 và Nhiệm vụ 2 được đưa ra nhằm thực hiện mục tiêu thứ nhất, và được chi tiết hóa trong Chương 3 với chủ đề "Tối ưu hiệu năng hệ thống truyền thông gói tin ngắn trong mạng chuyển tiếp hai chặng". Mục tiêu thứ hai được thực hiện qua Nhiệm vụ 3 và 4. Nhiệm vụ thứ 3 này được đặt trong phạm vi của Chương 4, với tiêu đề "Phân tích hiệu năng hệ thống truyền gói tin ngắn trong mạng chuyển tiếp sử dụng mặt phản xạ thông minh". Nhiệm vụ 4, với chủ đề "Thiết kế hệ thống truyền gói tin ngắn trong mạng đa truy nhập phi trực giao sử dụng mặt phản xạ thông minh", được trình bày trong Chương 5.

Chương 2. CƠ SỞ LÝ THUYẾT

2.1 Truyền thông tin gói ngắn

2.1.1 Cấu trúc gói tin ngắn

Hiện nay, thử thách lớn của hầu hết mạng di động thế hệ mới 5G và thế hệ tương lai 6G là truyền gói tin ngắn để đảm bảo yêu cầu uRLLCs. Gói tin ngắn đại diện cho phương thức trao đổi dữ liệu điển hình, được sinh ra từ các cảm biến và truyền tải trong các hệ thống giao tiếp máy với máy. Chi tiết cấu trúc của một gói tin ngắn được minh họa như trong Hình 2.1 [7].



Hình 2.1: Cấu trúc một gói tin ngắn.

Trong Hình 2.1, chúng ta có thể thấy cấu trúc của một gói dữ liệu ngắn, thường bao gồm metadata và dữ liệu. Cấu trúc của một metadata gồm phần màu đầu (P) và phần đầu (H). Metadata chứa các bits thông tin bổ sung nhằm thực hiện đúng nhiệm vụ của giao thức truyền không dây. Dữ liệu chứa các bit thông tin cần được truyền cho nguồn nhận. Khác biệt so với việc áp dụng gói tin dài trong các mạng di động trước đây (4G trở về trước), phương pháp truyền dữ liệu qua gói tin ngắn có những ưu điểm và khuyết điểm được mô tả dưới đây:

❖ Ưu điểm:

- Dữ liệu chứa trong gói tin ngắn có dung lượng nhỏ (khoảng .100bits ./dữ liệu truyền) [47]. Do đó, quá trình truyền gói tin ngắn có khả năng chống nhiễu tốt hơn khi truyền qua kênh truyền.
- Gói tin ngắn, hỗ trợ bởi mạng 5G và 6G, đủ khả năng đáp ứng các yêu cầu về độ trễ cực thấp trong hệ thống giao tiếp uRLLCs.

❖ Khuyết điểm:

- Phần metadata trong gói dữ liệu ngắn có kích thước bằng với phần dữ liệu nên việc mã hóa thông tin điều khiển không hiệu quả sẽ tác động đến hiệu năng của toàn bộ quá trình truyền gói.
- Cơ sở lý thuyết thông tin cho các mạng thông tin và di động hiện tại được phát triển dựa trên giả định sử dụng cho gói tin dài, do đó phần lớn không phù hợp cho gói tin ngắn.

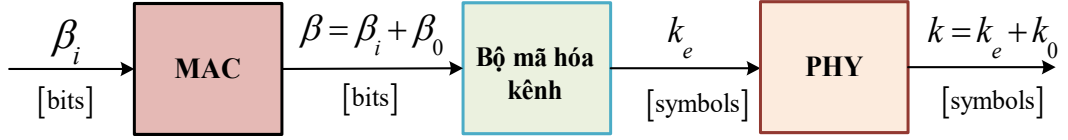
2.1.2 Đánh giá truyền thông tin gói ngắn

Cho đến nay, phần lớn các hệ thống không dây tiến tiến truyền dữ liệu dưới dạng gói tin. Mỗi gói tin được truyền trên kênh không dây không chỉ chứa thông tin (dưới dạng bit) dành cho thiết bị thu, mà còn bao gồm thêm các bit điều khiển cần thiết để bảo đảm hoạt động chính xác của các giao thức vô tuyến. Như trong Hình 2.2, một gói tin bao gồm β bit dữ liệu, được tạo thành từ β_i bits thông tin và β_0 bit bổ sung, chứa thông tin điều khiển từ lớp điều khiển truy nhập mạng (MAC) và các lớp cao hơn. Thông thường, các bit dữ liệu được mã hóa thành một khối k_e ký hiệu dữ liệu (số phức) để tăng cường độ tin cậy trong truyền dữ liệu gói tin. Cuối cùng, k_0 ký hiệu bổ sung được thêm vào để kích hoạt chức năng phát hiện gói tin, đồng bộ hóa hiệu suất (thời gian và tần số), hoặc ước lượng CSI, mà bộ thu cần để bù đắp cho sự méo mó của tín hiệu truyền đi khi qua kênh truyền. Do đó, chiều dài tổng cộng của gói tin là $k = k_e + k_0$. Hầu hết các hệ thống không

dây hiện nay thường có $\beta_i \gg \beta_0$ và $k_e \gg k_0$, vì vậy $\mathcal{R} = \frac{\beta_i}{k_e} = \frac{\beta}{k}$ là số bit thông tin

trên mỗi ký hiệu phức (hay số bit dữ liệu truyền tải mỗi giây trên mỗi đơn vị băng thông). Giá trị này đại diện cho tốc độ truyền mạng và cũng được xem là chỉ số đánh giá hiệu năng trong các hệ thống liên lạc.

Tiếp theo, Hình 2.3 mô hình truyền tin cơ bản theo lý thuyết thông tin.



Hình 2.2: Sơ đồ khối minh họa cách tạo một gói tin.



Hình 2.3: Mô hình khối của một hệ thống truyền tin.

Mô hình một hệ thống truyền tin cơ bản gồm bộ mã hóa, bộ giải mã và kênh truyền tin như được minh họa trong Hình 2.3. Về mặt toán học, bộ mã hoá được biểu diễn dưới dạng hàm f_β biến đổi β bit thông tin $A_1, \dots, A_\beta$ thành chuỗi ký tự $X_1, \dots, X_k$ phát trên kênh truyền. Chiều dài của chuỗi ký tự, k , định nghĩa chiều dài của gói tin (hay khối tin), chuỗi $X_1, \dots, X_k$ được gọi là từ mã. Các ký tự được truyền trên kênh truyền tuân theo điều kiện $\frac{1}{k} \sum_{i=1}^k |X_i|^2 \leq \rho$ với ρ là công suất giới hạn trung bình. Nhiệm vụ của bộ giải mã là đoán các bit thông tin $A_1, \dots, A_\beta$ từ các đầu ra k kênh $Y_1, \dots, Y_k$ sau khi qua kênh truyền. Tương tự, quá trình giải mã khái quát qua hàm g_k , ánh xạ tín hiệu đầu ra kênh $Y_1, \dots, Y_k$ thành chuỗi bit ước lượng $\hat{A}_1, \dots, \hat{A}_\beta$.

Định nghĩa ε là xác suất lỗi của gói tin, hay còn là xác suất mà bộ giải mã không chính xác chuỗi bit thông tin ban đầu $A_1, \dots, A_\beta$. Một mô hình truyền tin lý tưởng là thiết kế sao cho ε nhỏ và $\mathcal{R}$ lớn nhất có thể. Polyanskiy và các đồng nghiệp trong [9] đã chứng minh được tốc độ truyền cực đại (hay tốc độ mã hóa tối đa) với chiều dài $k \geq 100$ bits và tỉ lệ lỗi khối (BLER) ε có dạng như sau:

$$\mathcal{R}^*(k, \varepsilon) = C - \sqrt{\frac{V}{k}} Q^{-1}(\varepsilon) + O\left(\frac{\log k}{k}\right), \quad (2.1)$$

với $O\left(\log \frac{k}{k}\right)$ bao gồm số hạng còn lại của của bậc $\frac{\log k}{k}$ và có giá trị tùy vào kênh truyền; C là chỉ dung lượng kênh Shannon ; Q^{-1} đại diện cho hàm ngược của hàm Q-function [48]; V là độ phân tán kênh truyền [9]. Xét trường hợp kênh truyền nhiễu trắng AWGN, phương trình (2.1) được viết lại bằng cách thay số hạng $O\left(\log \frac{k}{k}\right) = \frac{\log k}{2k}$ như sau [49]:

$$\mathcal{R}^*(k, \varepsilon) = C(\gamma) - \sqrt{\frac{V(\gamma)}{k}} Q^{-1}(\varepsilon) + \frac{\log k}{2k}, \quad (2.2)$$

với $C(\gamma) = \log(1 + \gamma)$; $V(\gamma) = \left(1 - \frac{1}{(1 + \gamma)^2}\right) (\log_2 e)^2$, với γ là SNR nhận được tại nguồn thu nào đó. Từ (2.2), BLER trung bình được tính xấp xỉ theo $\mathcal{R}^*$ và k như sau [50]:

$$\bar{\varepsilon} \approx \mathbb{E} \left\{ Q \left(\frac{C(\gamma) - \mathcal{R}^*}{\sqrt{V(\gamma)/k}} \right) \right\}. \quad (2.3)$$

Chúng ta thấy rằng dạng đóng cho BLER trong (2.3) khó có thể tính được do hàm $Q\left(\frac{C(\gamma) - \mathcal{R}^*}{\sqrt{V(\gamma)/k}}\right)$ khá phức tạp. Do vậy, chúng ta sử dụng xấp xỉ của hàm $Q(\cdot) \approx \Psi(\gamma)$ như trong [51. CT. (14)], với

$$\Psi(\gamma) = \begin{cases} 1, & \gamma \leq \rho_L, \\ \frac{1}{2} - v\sqrt{k}(\gamma - \theta), & \rho_L < \gamma < \rho_H, \\ 0, & \gamma \geq \rho_H, \end{cases} \quad (2.4)$$

$$v = \frac{1}{2\pi\sqrt{2^{2\mathcal{R}^*} - 1}}, \quad (2.5)$$

$$\theta \triangleq 2^{\mathcal{R}^*} - 1, \quad (2.6)$$

$$\rho_H = \theta + \frac{1}{2v\sqrt{k}}, \quad (2.7)$$

$$\text{và} \quad \rho_L = \theta - \frac{1}{2v\sqrt{k}}. \quad (2.8)$$

Bằng cách thay (2.4) vào (2.3), BLER trung bình được viết lại như sau:

$$\begin{aligned} \bar{\varepsilon} &\approx \int_0^\infty \Psi(\gamma) f_\gamma(\gamma) d\gamma \\ &= v\sqrt{k} \int_{\rho_L}^{\rho_H} F_\gamma(\gamma) d\gamma. \end{aligned} \quad (2.9)$$

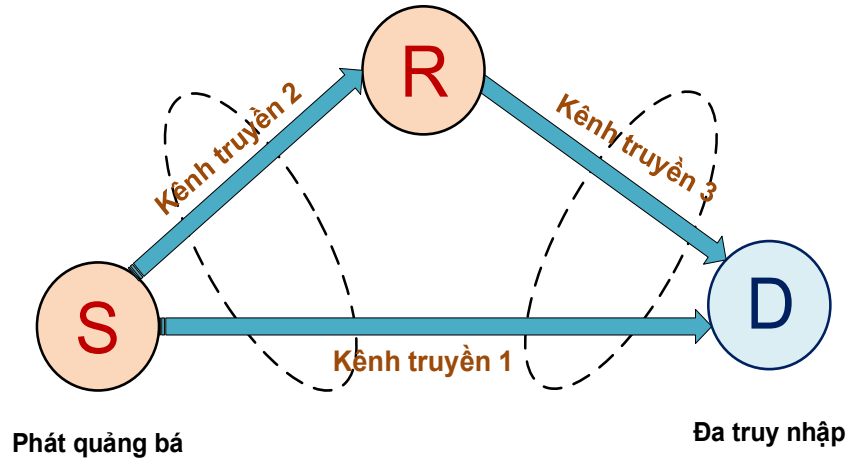
Hiện nay, BLER trung bình trong phương trình (2.9) được coi là một chỉ số đánh giá hiệu năng quan trọng cho các hệ thống SPC.

2.2 Truyền thông cộng tác

Cho đến nay, truyền thông cộng tác được khẳng định là một phương pháp tiềm năng, có khả năng tăng cường hiệu năng của hệ thống liên lạc không dây, bằng cách tối ưu dung lượng, cải thiện độ phân tập và mở rộng phạm vi phủ sóng, qua sự hợp tác giữa các nút chuyển tiếp [14, 15, 52, 53]. Hiện nay, truyền thông cộng tác đang nhận được ưu tiên đặc biệt trong việc áp dụng cho mạng di động 5G và mạng tương lai, với mục tiêu đáp ứng các yêu cầu khắc khe về sự tin cậy và thời gian trễ của uRLLCs [17, 54]. Ý tưởng chính trong truyền thông cộng tác là các đầu cuối mạng kết hợp an-ten của mình để tạo ra hệ thống an-ten ảo [14]. Điều này giúp mạng cộng tác nâng cao hiệu quả phân tập mà không cần trang bị thêm nhiều an-ten ở các điểm cuối mạng. Hình 2.4 minh họa một mô hình kênh cơ bản trong truyền thông cộng tác.

Trong Hình 2.4, quá trình truyền dữ liệu từ nút nguồn (S) đến nút đích (D) được hỗ trợ bởi nút chuyển tiếp (R) và diễn ra qua hai giai đoạn: i) Giai đoạn quảng bá, S gửi thông tin đến R và D; ii) Giai đoạn truy nhập, S và R cùng truyền thông tin

nhưng chỉ D nhận. Nhờ vậy, nút D có khả năng thu nhận tín hiệu từ cả S và R, làm tăng độ tin cậy trong việc giải mã tín hiệu tại D. Đồng thời, hệ thống này cũng cải thiện sự phân tập không gian do khả năng tín hiệu từ S đến D theo nhiều đường khác nhau, từ đó nâng cao hiệu quả của hệ thống [15].



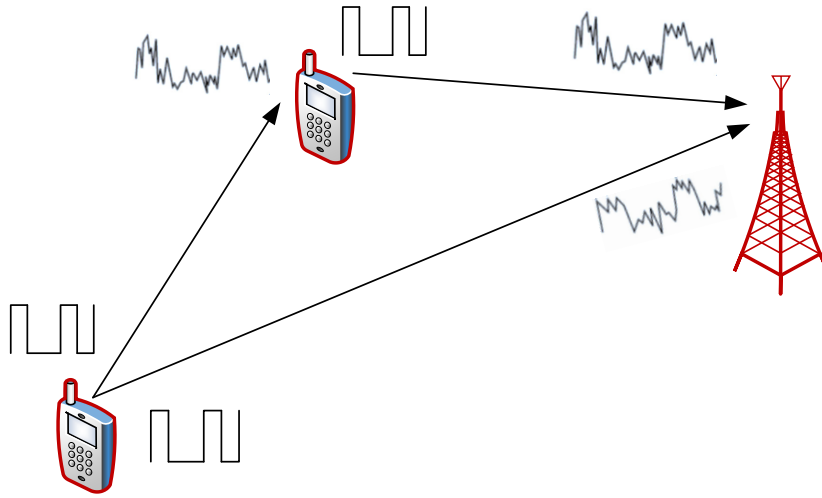
Hình 2.4: Mô hình kênh của truyền thông cộng tác.

Để thực hiện vai trò truyền tín hiệu từ S đến D, nút R áp dụng cả hai phương pháp xử lý tín hiệu: giải mã và chuyển tiếp DF [55] hoặc khuếch đại và chuyển tiếp AF [56].

Đầu tiên, kỹ thuật DF được thực hiện như được minh họa trong Hình 2.5, ở đây nút R đóng vai trò của một bộ lặp mạng, giải mã tín hiệu nhận được từ nút S trong giai đoạn truyền đầu tiên. Hơn nữa, kỹ thuật DF bao gồm hai loại chuyển tiếp tương ứng với hai khả năng xảy ra ở nút R là chuyển tiếp có chọn lựa (SDF) và chuyển tiếp DF cố định (FDF). Ở SDF, nút R chỉ chuyển tiếp tín hiệu đến nút D khi R giải mã đúng tín hiệu mà nút S muốn truyền đến D. Ngược lại, trong FDF nút R chuyển tiếp tín hiệu đến D mà không cần xác định tín hiệu được giải mã chính xác hay không. Đặc biệt, cả hai quá trình chuyển tiếp FD sẽ loại trừ được sự ảnh hưởng của nhiễu [57, 58].

Trong kỹ thuật chuyển tiếp AF, R khuếch đại tín hiệu nhận được từ nút S và sau đó chuyển tiếp toàn bộ tín hiệu này về nút D, như trình bày trong Hình 2.6. So với

DF, kỹ thuật AF không phải thực hiện quá trình giải điều chế tín hiệu tại R nhưng nhược điểm là chuyển tiếp và khuếch đại nhiễu.



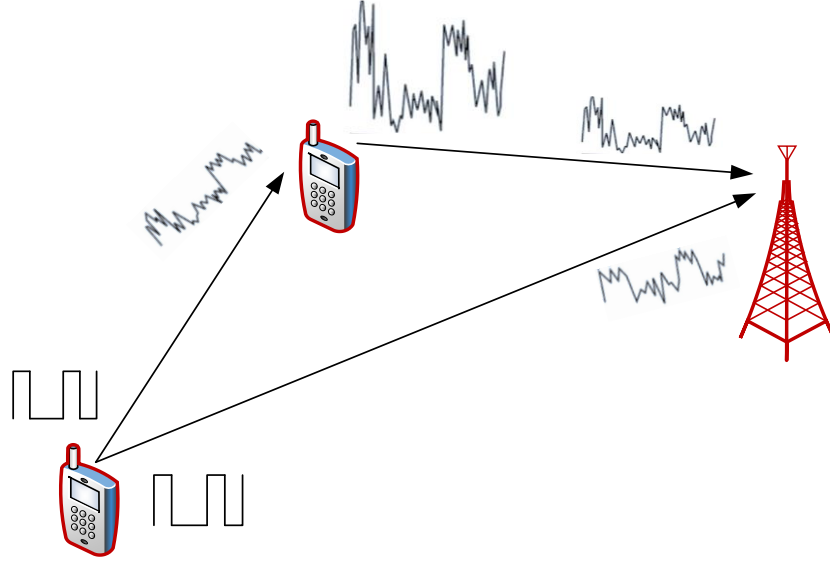
Hình 2.5: Kỹ thuật chuyển tiếp DF.

Để tối ưu khả năng phân tập không gian trong hệ thống truyền thông công tác, nút S có thể được thiết kế với đa an-ten, sử dụng kỹ thuật lựa chọn an-ten phát (TAS). Các nút chuyển tiếp có thể được thiết kế đơn an-ten và áp dụng phương pháp lựa chọn nút chuyển tiếp, trong khi nút D sử dụng các kỹ thuật phân tập kết hợp như kết hợp lựa chọn (SC), cân bằng độ lợi (EGC) và tỉ số cực đại (MRC).

❖ Kỹ thuật chọn lựa an-ten phát TAS:

Đa dạng an-ten là kỹ thuật hiệu quả để giảm suy giảm kênh truyền thông không dây, như hiện tượng đa đường và bóng râm [59, 60]. Tuy nhiên, mỗi an-ten thêm vào sẽ kèm theo chuỗi tần số vô tuyến và tăng tiêu thụ năng lượng, làm tăng chi phí và độ phức tạp. Việc chọn lựa an-ten giúp giảm chi phí và phức tạp, đồng thời cải thiện hiệu suất hệ thống đáng kể [4], [5]. Chọn an-ten có thể thực hiện tại nguồn phát dưới dạng chọn lựa an-ten phát TAS. Giả sử nguồn phát có N_T an-ten và nguồn nhận có thông tin trạng thái kênh hoàn hảo. Dựa trên thông tin phản hồi CSI một phần [61] từ nguồn nhận, nguồn phát xác định chỉ số an-ten tối ưu để tối đa SNR nhận được. Cuối cùng, nguồn phát chỉ sử dụng an-ten tối ưu trong số N_T an-ten để truyền dữ liệu. SNR của an-ten được chọn có thể biểu diễn như sau:

$$\gamma^{\text{TAS}} = \max_{i=1,2,\dots,N_T} \gamma_i. \quad (2.10)$$



Hình 2.6: Kỹ thuật chuyển tiếp AF.

❖ Kỹ thuật lựa chọn nút chuyển tiếp bán phần (PRS):

Trong các hệ thống không dây hiện đại như mạng di động và mạng cảm biến, việc triển khai thường gặp phải những thách thức do hạn chế kích thước và khó khăn trong tích hợp nhiều an-ten trên các nút không dây. Vì vậy, các nút thường chỉ được trang bị một an-ten đơn và hoạt động ở chế độ bán song công. Trong hệ thống truyền thông hợp tác sử dụng N nút chuyển tiếp $(R_1, R_2, \dots, R_N)$, các giao thức hợp tác đa nút đạt được độ lợi đa dạng $N+1$ [62, 63]. Việc tăng số lượng nút chuyển tiếp giúp cải thiện đáng kể hiệu quả truyền dữ liệu. Tuy nhiên, chế độ bán song công yêu cầu $N+1$ khe thời gian để nguồn và các nút chuyển tiếp truy cập kênh, làm giảm hiệu suất sử dụng kênh xuống còn $1/(N+1)$. Để giải quyết vấn đề này, chiến lược chọn nút chuyển tiếp tốt nhất R_b đã được đề xuất [64]. Công thức xác định R_b như sau:

$$R_b : \min(\Omega_{S,R_b}, \gamma_{R_b,D}) = \max_{j=1,2,\dots,N} \left(\min(\gamma_{S,R_j}, \gamma_{R_j,D}) \right), \quad (2.11)$$

Trong đó, γ_{S,R_j} và $\gamma_{R_j,D}$ lần lượt là SNR tức thời của các liên kết $S \rightarrow R_j$ và $R_j \rightarrow D$. Chiến lược này có thể được thực hiện phân tán, với mỗi nút R_j đặt một bộ định thời: $T_j = \frac{C}{\min(\gamma_{S,R_j}, \gamma_{R_j,D})}$, với C là hằng số. Nút có bộ định thời hết hạn

sớm nhất sẽ được chọn làm nút chuyển tiếp tối ưu. Một trong những phương pháp đơn giản và hiệu quả nhất là lựa chọn nút chuyển tiếp bán phần, trong đó việc xác định nút chuyển tiếp tốt nhất chỉ dựa trên thông tin trạng thái CSI của liên kết giữa nguồn và nút chuyển tiếp, được mô tả như sau:

$$R_b : \gamma_{S,R_b} = \max_{i=1,2,\dots,N} (\gamma_{S,R_i}), \quad (2.12)$$

Phương pháp lựa chọn nút chuyển tiếp bán phần có nhiều ưu điểm nổi trội như

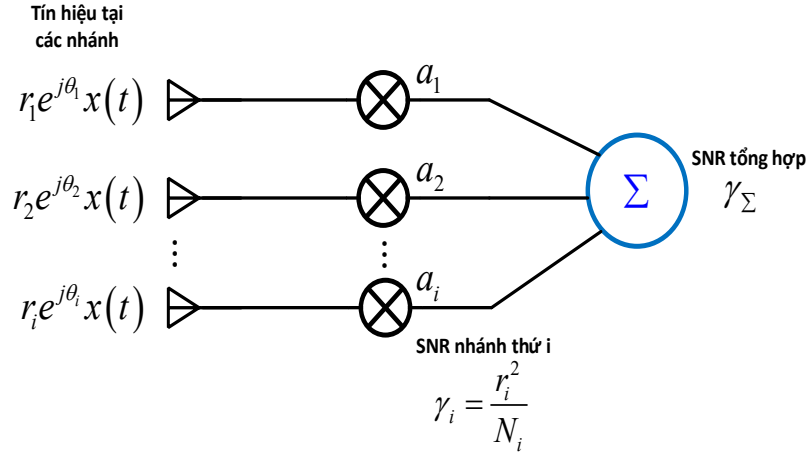
- Giảm yêu cầu tính toán bằng cách chỉ sử dụng CSI của liên kết nguồn - nút chuyển tiếp.
- Tăng hiệu suất sử dụng kênh, giảm tín hiệu điều khiển và tránh va chạm.
- Kéo dài tuổi thọ mạng nhờ chọn nút chuyển tiếp tối ưu.
- Có thể ứng dụng cho mạng thế hệ mới như 5G và 6G nhờ tính phân tán và yêu cầu thông tin tối thiểu.

❖ Tại nút D áp dụng các phương pháp phân tập kết hợp SC, cân bằng độ lợi EGC, và tỉ số cực đại MRC. Các phương pháp phân tập này hoạt động dựa trên nguyên tắc kết hợp tuyến tính các đầu ra từ bộ kết hợp, được minh họa chi tiết trong Hình 2.7. Giả sử có M nhánh tín hiệu độc lập, tín hiệu tại mỗi nhánh được biểu diễn dưới dạng phức . . với $i = 1, \dots, M$, trong đó: r_i là biên độ của tín hiệu ở nhánh thứ i ; θ_i là pha của tín hiệu nhánh i ; $x(t)$ là tín hiệu gốc. Mỗi tín hiệu nhánh sau đó được nhân với một trọng số phức a_i , được định nghĩa $a_i = e^{-j\theta_i}$. Điều này giúp loại bỏ pha giúp loại bỏ pha θ_i , chỉ còn lại biên độ của tín hiệu. Sau đó, tất cả các tín hiệu đã cân pha được cộng lại để tạo ra tín hiệu đầu ra duy nhất. Các tín hiệu đã nhân trọng số được đưa vào bộ cộng (ký hiệu Σ) để tạo ra tín hiệu

đầu ra duy nhất. Khi tín hiệu từ tất cả các nhánh được cộng lại, giá trị SNR tổng hợp sẽ là:

$$\gamma_{\Sigma} = \sum_{i=1}^M \gamma_i, \quad (2.13)$$

với γ_i là SNR của từng nhánh tín hiệu.



Hình 2.7: Bộ kết hợp tuyến tính.

- Phương pháp lựa chọn nhánh SC: trong điều kiện kênh Rayleigh fading và nhiễu Gaussian trắng cộng (AWGN), hệ thống sẽ chọn nhánh có giá trị biên độ lớn nhất: $A_{\max} = \max(A_i)$ [65, 66]. SNR trung bình của đầu ra bộ kết hợp được tính bằng:

$$\bar{\gamma}_{SC} = \gamma \sum_{i=1}^M \frac{1}{i}. \quad (2.14)$$

Công thức (2.14) này cho thấy rằng tổng hợp càng nhiều nhánh sẽ cải thiện SNR, nhưng lợi ích sẽ giảm dần.

- Phương pháp lựa chọn nhánh MRC: kết hợp tín hiệu từ nhiều nhánh anten với các trọng số tối ưu [67, 68]. Tín hiệu đầu ra là tổng có trọng số của các nhánh, được biểu diễn bằng: $r = \sum_{i=1}^M a_i r_i$. Nếu mỗi nhánh có mật độ công suất nhiễu là

$\frac{N_0}{2}$, tổng mật độ công suất nhiễu đầu ra là $\frac{N_{\text{tong}}}{2} = \sum_{i=1}^M a_i^2 N_0 / 2$. Khi đó, SNR đầu ra được tính bởi:

$$\gamma_{\Sigma} = \frac{r^2}{N_{\text{tong}}} = \frac{1}{N_0} \frac{\left(\sum_{i=1}^M a_i r_i \right)^2}{\sum_{i=1}^M a_i^2}. \quad (2.15)$$

Để tối đa γ_{Σ} , ta cần chọn các trọng số a_i sao cho ưu tiên các nhánh có SNR cao hơn. Điều này đạt được bằng cách đặt $a_i^2 \propto \frac{r_i^2}{N_0}$. Với cách làm này: $a_i^2 = \frac{r_i^2}{N_0}$

. Khi đó, SNR sau khi kết hợp là: $\gamma_{\Sigma} = \frac{r^2}{N_0} = \sum_{i=1}^M \gamma_i$. Vậy SNR trung bình đầu ra cũng được xác định bởi:

$$\bar{\gamma}_{MRC} = M\gamma. \quad (2.16)$$

- Phương pháp kết hợp cân bằng độ lợi EGC: để áp dụng MRC, cần biết SNR tức thời, điều này khó xác định. Trong kỹ thuật đơn giản hơn EGC, trọng số a_i được cố định dưới dạng $a_i = e^{-j\theta_i}$ [69]. Giả sử rằng mật độ công suất nhiễu đều bằng $\frac{N_0}{2}$ ở mỗi nhánh, SNR của đầu ra bộ kết hợp được cho bởi:

$$\gamma_{\Sigma} = \frac{1}{N_0 M} \left(\sum_{i=1}^M r_i \right)^2. \quad (2.17)$$

Dựa vào (2.14), (2.15), (2.17) ta thấy được phương pháp MRC là tối ưu nhất về mặt hiệu suất SNR đầu ra vì MRC tận dụng tất cả các nhánh một cách hiệu quả bằng cách tối đa SNR của tín hiệu tổng hợp. Trong khi đó, SC chỉ sử dụng nhánh có SNR tốt nhất và EGC không cân nhắc trọng số theo SNR, khiến cho MRC vượt trội hơn trong môi trường có nhiễu và fading.

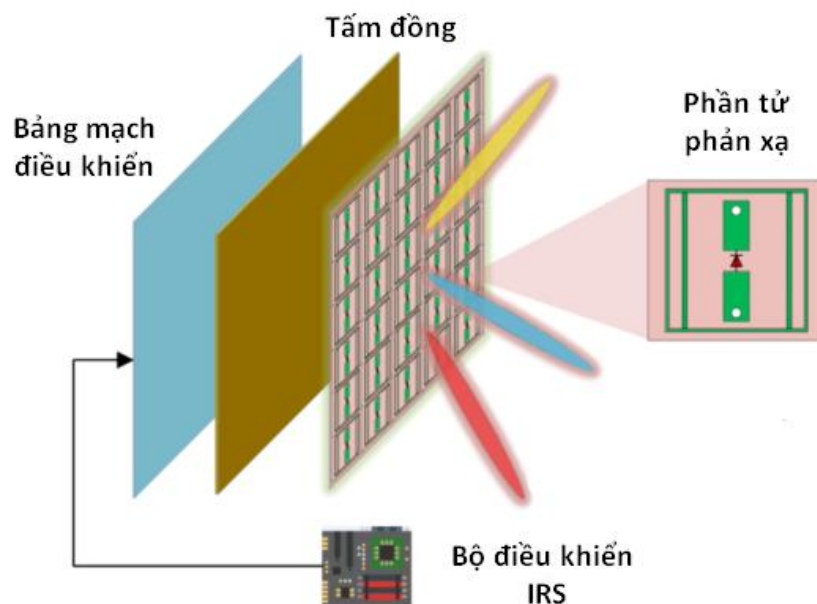
Nói chung, truyền thông cộng tác là được coi là một phương pháp hứa hẹn, mang lại nhiều lợi ích đáng kể như độ phân tán không gian cao, mở rộng vùng phủ sóng, tiết kiệm năng lượng và nâng cao hiệu năng, không chỉ dành cho các hệ thống

mạng không dây hiện tại. Điều này cũng giải thích được rằng cộng đồng nghiên cứu hiện nay đang tập trung vào việc tích hợp SPC vào truyền thông cộng tác, để đáp ứng yêu cầu uRLLCs cho mạng di động 5G và thậm chí tương lai 6G [11, 22].

2.3 Mặt phản xạ thông minh

2.3.1 Cấu trúc IRS thụ động

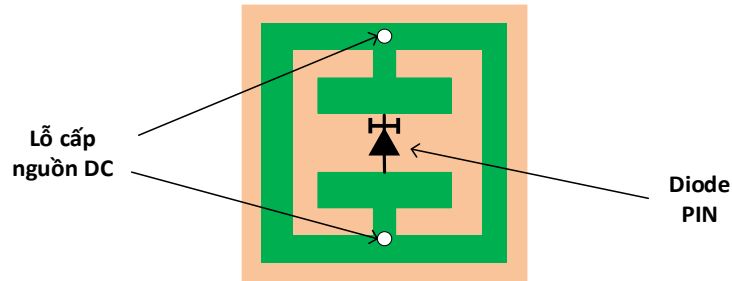
Cấu trúc thụ động của một mặt phản xạ IRS như được minh họa Hình 2.8, được chia thành ba lớp chính và một thiết bị điều khiển tự động [26]. Lớp ngoài cùng gồm nhiều phần tử phản xạ được in trên một tấm điện môi, có chức năng phản xạ trực tiếp các tín hiệu đến. Lớp giữa là một tấm đồng, đóng vai trò quan trọng trong việc giảm thiểu tiêu hao năng lượng trong quá trình phản xạ của IRS. Lớp trong cùng chứa bảng mạch điều khiển, có nhiệm vụ kích thích các phần tử phản xạ và điều chỉnh biên độ, pha phản xạ theo thời gian thực. Lớp này được điều khiển bởi một bộ điều khiển thông minh, chịu trách nhiệm giao tiếp với các thành phần khác trong mạng (như trạm gốc/điểm truy cập và người dùng) thông qua một kênh truyền riêng (có thể là dây dẫn hoặc kênh điều khiển vô tuyến) để ra quyết định điều chỉnh phù hợp các phần tử phản xạ.



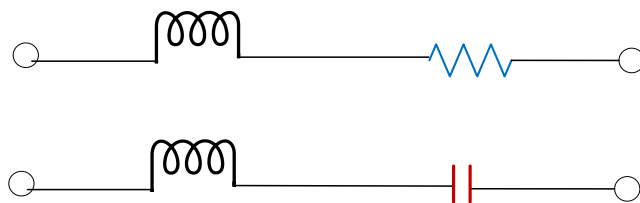
Hình 2.8: Cấu trúc của mặt phản xạ thông minh IRS [27].

Ngoài ra, IRS còn có thể tích hợp thêm các cảm biến trong lớp ngoài cùng, xen kẽ với các phần tử phản xạ, giúp cảm nhận các tín hiệu vô tuyến xung quanh nhằm hỗ trợ bộ điều khiển thông minh tối ưu hệ số phản xạ.

❖ Cấu hình các phần tử IRS: hiện có ba phương pháp chính để cấu hình phần tử IRS nhằm đạt được phản xạ mong muốn: 1) điều khiển cơ học (qua xoay và dịch chuyển cơ học); 2) sử dụng vật liệu đặc biệt (như tinh thể lỏng và graphene); và 3) thiết bị điện tử (như diode PIN, transistor, hoặc công tắc vi cơ điện tử) [51]. Phương pháp dùng thiết bị điện tử được áp dụng rộng rãi nhờ thời gian phản hồi nhanh, tổn thất phản xạ thấp, tiêu thụ năng lượng và chi phí phần cứng tương đối thấp. Ví dụ, diode PIN có thể chuyển giữa trạng thái “BẬT” và “TẮT” khi cấp điện áp phân cực, cho phép tạo ra sự khác biệt pha là π cho tín hiệu tới, như minh họa trong Hình 2.9. Tần số chuyển mạch của diode PIN có thể lên đến 5 MHz, tương đương thời gian chuyển mạch 0,2 μ s, phù hợp cho các ứng dụng di động với kênh thay đổi nhanh.



a) Phần tử phản xạ



b) Mạch tương đương của Diode PIN

Hình 2.9: Điều chỉnh phần tử phản xạ dựa trên PIN diode.

Ngoài điều chỉnh pha, kiểm soát biên độ phản xạ của mỗi phần tử IRS giúp tăng linh hoạt trong việc định hình tín hiệu để đạt được các mục tiêu truyền thông khác nhau. Một cách phổ biến để điều chỉnh biên độ là thay đổi điện trở trong mỗi phần tử phản xạ [70], cho phép kiểm soát năng lượng tín hiệu tới và đạt dải động của biên độ phản xạ trong khoảng $[0, 1]$.

2.3.2 Mô hình tín hiệu và kênh qua IRS

Mô hình hệ thống thông tin cơ bản từ điểm-đến-điểm được minh họa như trong Hình 2.10. Tầm IRS thụ động được giả định có N yếu tố phản xạ, hỗ trợ truyền thông tin từ nguồn phát đơn an-ten đến nguồn thu đơn an-ten, và hệ thống thông tin là băng hẹp ở một tần số sóng mang nhất định. Tần số sóng mang và băng thông hệ thống lần lượt được ký hiệu là f_c và B theo đơn vị hertz (Hz), với $B \ll f_c$. Để đơn giản, Luận án tập trung phân tích một đường tín hiệu truyền từ máy phát đến máy thu thông qua sự phản xạ trên một phần tử phản xạ n của IRS với $n = 1, 2, \dots, N$.

Giai đoạn 1: nguồn phát S truyền tín hiệu đến yếu tố phản xạ thứ n của IRS. Tín hiệu nhận được tại n trên IRS là:

$$y_{in,n}(t) = \text{Re}\left\{\alpha_{1,n}e^{-j\zeta_{1,n}}x(t)e^{j2\pi ft_c}\right\}, \quad (2.18)$$

với $\alpha_{1,n}e^{-j\zeta_{1,n}}$ là hệ số kênh truyền phức băng cơ sở từ nguồn phát đến phần tử n , trong đó $\alpha_{1,n}$ và $\zeta_{1,n}$ lần lượt là biên độ và độ dịch pha của kênh tần số phẳng trong hệ thống băng hẹp của yếu tố phản xạ n trong IRS. Đặt $\beta_n \in [0, 1]$ và $t_n \in [0, 1/f_c]$ là độ suy hao biên độ và độ trễ thời gian do phần tử phản xạ thứ n gây ra. Nếu bỏ qua nhiễu, tín hiệu phản xạ bởi phần tử n được biểu diễn như sau:

$$\begin{aligned}
y_{out,n}(t) &= \beta_n y_{in,n}(t - t_n) = \text{Re} \left\{ \beta_n \alpha_{1,n} e^{-j\zeta_{1,n}} x(t - t_n) e^{j2\pi f_c(t - t_n)} \right\} \\
&\approx \text{Re} \left\{ \beta_n e^{j\theta_n} \alpha_{1,n} e^{-j\zeta_{1,n}} x(t) e^{j2\pi f_c(t)} \right\},
\end{aligned} \tag{2.19}$$

với $x(t - t_n) \approx x(t)$ do $t_n \leq 1/f_c \ll 1/B$, và $\theta_n \triangleq -j2\pi f_c t_n \in [0, 2\pi)$ là sự dịch pha do phần tử phản xạ thứ n gây ra. Ngoài ra, đặt tín hiệu băng cơ sở của $y_{in,n}(t)$ và $y_{out,n}(t)$ là $x_{in,n}(t) \triangleq \alpha_{1,n} e^{-j\zeta_{1,n}} x(t)$ và $x_{out,n}(t) \triangleq \beta_n e^{j\theta_n} \alpha_{1,n} e^{-j\zeta_{1,n}} x(t) e^{j2\pi f_c(t)}$, tương ứng. Góc lệch pha $\theta_n \in [0, 2\pi]$; h_n và g_n lần lượt là hệ số kênh truyền từ nút S đến IRS và từ IRS đến nút D . Vậy mô hình tín hiệu phản xạ (đầu ra) từ phần tử n của IRS trong băng tần cơ sở được biểu diễn như sau:

$$x_{out,n}(t) = \beta_n e^{j\theta_n} x_{in,n}(t). \tag{2.20}$$

Giai đoạn 2: tín hiệu truyền từ yếu tố phản xạ thứ n của IRS đến máy thu D . Tương tự như trên, tín hiệu nhận được tại D được biểu diễn như sau:

$$y_{r,n}(t) = \text{Re} \left\{ \left[\alpha_{1,n} e^{-j\zeta_{1,n}} \beta_n e^{j\theta_n} \alpha_{2,n} e^{-j\zeta_{2,n}} x(t) \right] e^{j2\pi f_c t} \right\}, \tag{2.21}$$

với $\alpha_{2,n} e^{-j\zeta_{2,n}}$ là kênh truyền tần số phẳng băng hẹp tương đương của chặng từ IRS đến D . Đặt $h_n \triangleq \alpha_{1,n} e^{-j\zeta_{1,n}}$ và $g_n = \alpha_{2,n} e^{-j\zeta_{2,n}}$, mô hình kênh liên tiếp từ nguồn phát S đến nguồn thu D qua phần tử IRS được cho như sau:

$$y_n(t) = \beta_n e^{j\theta_n} h_n g_n x(t). \tag{2.22}$$

Như vậy, kênh truyền nối tiếp từ nguồn phát đến IRS đến nguồn thu là tích của ba thành phần: kênh truyền từ nguồn phát đến phần tử phản xạ thứ n (h_n); hệ số phản xạ IRS $\beta_n e^{j\theta_n}$; và kênh truyền từ IRS đến máy thu (g_n).

Mở rộng phân tích với IRS bao gồm N phần tử và các phần tử phản xạ tín hiệu một cách độc lập. Khi đó tín hiệu tổng cộng nhận được tại nguồn thu D là:

$$y_D(t) = \left(\sum_{n=1}^N h_n g_n \beta_n e^{j\theta_n} \right) x(t). \tag{2.23}$$

Từ (2.23), ta có SNR tức thời nhận được tại nguồn thu D như sau:

$$\gamma_D \sim \left| \sum_{n=1}^N h_n g_n \beta_n e^{j\theta_n} \right|^2 = \left| \sum_{n=1}^N \alpha_{1,n} \alpha_{2,n} \beta_n e^{j(\theta_n - \zeta_{1,n} - \zeta_{2,n})} \right|^2. \quad (2.24)$$

Để tối ưu SNR, các pha của tín hiệu phải được căn chỉnh sao cho đồng pha. Điều này yêu cầu pha trong số mũ phức phải thỏa mãn $\theta_n - \zeta_{1,n} - \zeta_{2,n} = \text{const}$ (hằng số). Do đó, pha của tín hiệu tại phần tử thứ n của IRS nên được điều chỉnh [71]:

$$\theta_n = \zeta_{1,n} + \zeta_{2,n} + \phi, \quad (2.25)$$

với ϕ là một hằng số pha chung (có thể chọn $\phi = 0$ để đơn giản). Việc điều chỉnh này đảm bảo tất cả các tín hiệu cộng hưởng cùng pha, dẫn đến SNR đạt cực đại.

❖ Các hệ số kênh h_n và g_n thường phụ thuộc vào suy hao đường truyền liên quan đến khoảng cách, bóng che quy mô lớn và fading đa đường quy mô nhỏ. Nếu n của IRS đủ xa cả nguồn phát và nguồn thu, với khoảng cách lần lượt $d_{1,n} = d_1$

và $d_{2,n} = d_2$ như trong Hình 2.10, thì $|h_n|^2 \propto c_1 \left(\frac{d_1}{d_0} \right)^{-\eta_1}$ và $|g_n|^2 \propto c_2 \left(\frac{d_2}{d_0} \right)^{-\eta_2}$ với

mọi n , trong đó c_1 (hoặc c_2) là suy hao đường truyền tại khoảng cách tham chiếu d_0 ; η là hệ số suy hao đường truyền tương ứng. Công suất tín hiệu trung bình nhận được thông qua phản xạ bởi phần tử n của IRS như sau:

$$P_n \propto \frac{1}{d_1^{\eta_1} d_2^{\eta_2}}. \quad (2.26)$$

Từ (2.26) ta thấy được kênh phản xạ qua phần tử n của IRS chịu suy hao đường truyền kép. Do đó, thực tế thiết kế một số lượng lớn các phần tử phản xạ IRS để bù đắp cho sự suy hao công suất do suy giảm kép là cần thiết, tức thiết kế đồng thời biên độ và/hoặc pha phản xạ của các hệ số kênh để đạt được độ lợi tạo chùm thụ động cao. Tuy nhiên, việc sử dụng bao nhiêu phần tử phản xạ, và việc tối ưu pha của các phần tử này như thế nào thì vẫn đang là vấn đề mở và còn tùy thuộc vào từng hệ thống thông tin khác nhau [72, 73].

❖ Nếu mặt IRS vừa phản xạ và truyền qua đồng thời (STAR-IRS), tức tín hiệu đến yếu tố thứ n của STAR-IRS được phân chia thành tín hiệu truyền qua và một tín hiệu phản xạ. Cả hai tín hiệu truyền qua và phản xạ đều được mô tả toán học như sau:

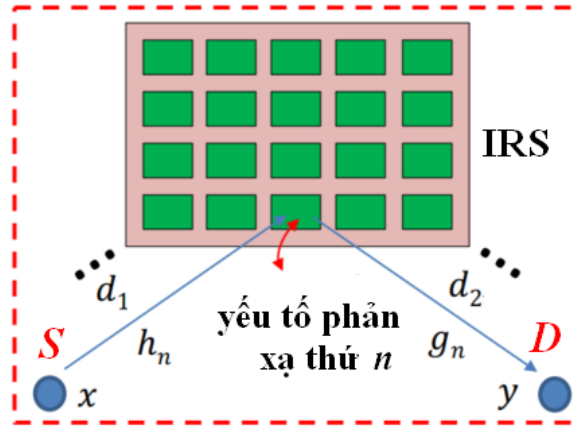
$$x_n^r(t) = \sqrt{\beta_n^r} e^{j\theta_n^r} x(t), \quad (2.27)$$

$$x_n^t(t) = \sqrt{\beta_n^t} e^{j\theta_n^t} x(t), \quad (2.28)$$

với $\sqrt{\beta_n^r} \in [0,1]$, $\theta_n^r \in [0,2\pi)$ và $\sqrt{\beta_n^t} \in [0,1]$, $\theta_n^t \in [0,2\pi)$ là các biên độ và góc lệch pha thứ n của hệ số phản xạ và truyền qua trong STAR-IRS, tương ứng. Góc lệch pha thứ n của hệ số phản xạ và truyền qua θ_n^r , θ_n^t có thể được chọn tách biệt với nhau. Tuy nhiên, các biên độ $\sqrt{\beta_n^r}$ và $\sqrt{\beta_n^t}$ của hệ số phản xạ và truyền qua đều tuân thủ định luật bảo toàn năng lượng. Nghĩa là, đối với mỗi yếu tố thứ n , tổng năng lượng của tín hiệu truyền qua và tín hiệu phản xạ phải bằng năng lượng của tín hiệu đến, tức là $|x_n^r(t)|^2 + |x_n^t(t)|^2 = |x(t)|^2$. Do đó, điều kiện cho độ lớn biên độ của hệ số truyền qua và phản xạ của mỗi yếu tố thứ n như sau:

$$\beta_n^r + \beta_n^t = 1, \quad \forall n \in N. \quad (2.29)$$

❖ Hiệu năng của hệ thống vô tuyến tích hợp IRS được cải thiện đáng kể khi trạm gốc (BS) và người dùng có đầy đủ thông tin trạng thái kênh CSI. Tuy nhiên, việc ước lượng kênh trong hệ thống này gặp nhiều thách thức do: 1) Số lượng lớn phần tử phản xạ trên IRS làm tăng đáng kể số lượng hệ số kênh và chuỗi thăm dò cần thiết; 2) Các phần tử phản xạ thụ động, chi phí thấp không được trang bị chuỗi tần số vô tuyến (RF) để hỗ trợ phát tín hiệu huấn luyện. Hiện nay, có hai phương pháp chính để ước lượng kênh trong các hệ thống IRS: dành cho IRS bán thụ động (có tích hợp cảm biến) và IRS thụ động hoàn toàn.

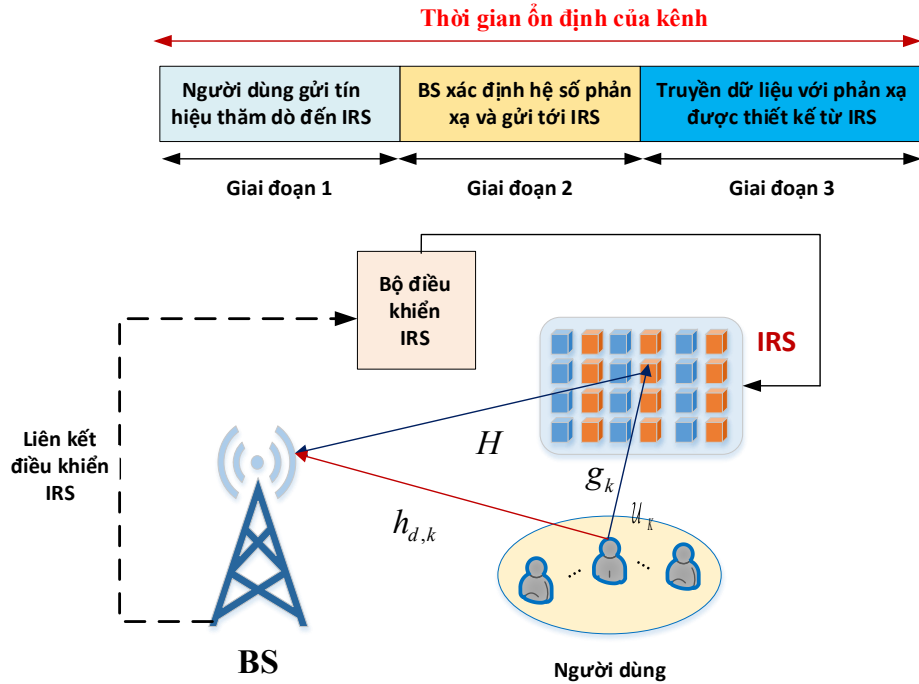


Hình 2.10: Mô hình tín hiệu qua IRS.

Để minh họa rõ hơn quy trình ước lượng kênh truyền cho IRS, chúng tôi trình bày giao thức truyền kênh uplink (từ người dùng đến trạm gốc) với IRS thụ động hoàn toàn. Trong đó, mỗi khoảng thời gian ổn định của kênh được chia thành ba pha liên tiếp, như mô tả trong Hình 2.11. Đầu tiên, người dùng truyền các tín hiệu thăm dò trực giao (pilots) đến trạm gốc BS, trong khi đó IRS thay đổi các hệ số phản xạ của mình theo một mẫu phản xạ được thiết kế trước. Dựa trên đó, trạm gốc có thể ước lượng cả kênh truyền trực tiếp từ người dùng đến BS và kênh truyền từ người dùng đến BS qua IRS (người dùng-IRS-BS). Thứ hai, dựa trên thông tin kênh CSI đã ước lượng, các hệ số phản xạ của IRS phục vụ cho truyền dữ liệu sẽ được thiết kế tại trạm gốc cùng với việc tạo chùm tia thu của BS và sau đó được gửi đến bộ điều khiển IRS qua liên kết backhaul. Thứ ba, bộ điều khiển IRS thiết lập các hệ số phản xạ theo các giá trị đã tính toán nhằm hỗ trợ truyền dữ liệu độc lập từ người dùng đến trạm gốc [74].

Ước lượng kênh với IRS thụ động gặp thách thức lớn khi phải đồng thời thiết kế chuỗi tín hiệu thăm dò, mẫu phản xạ IRS và thuật toán xử lý tín hiệu tại bộ thu để ước lượng chính xác kênh trực tiếp (người dùng-BS) và kênh qua IRS (người dùng-IRS-BS) với chi phí huấn luyện tối thiểu. Đây là vấn đề mới và phức tạp, không có trong các hệ thống không dây truyền thống. Để khắc phục, các phương pháp như mẫu phản xạ tất cả-ON và gom nhóm phân tử IRS đã được đề xuất. Bằng cách chia IRS thành các nhóm nhỏ dựa trên tính tương quan không gian của kênh,

chỉ cần ước lượng một kênh đại diện cho mỗi nhóm, sau đó nội suy để xác định các kênh còn lại.



Hình 2.11: Ước lượng kênh truyền trong mô hình IRS thụ động hoàn toàn (đường lên).

2.3.3 Ưu điểm của IRS

❖ Về bản chất, mặt IRS thực hiện chức năng chuyển tiếp tín hiệu từ S đến D tương tự như vai trò của một nút chuyển tiếp. Tuy nhiên, khả năng hoạt động của mặt IRS không nhất thiết phải cố định trong các khoảng thời gian, như hoạt động của nút chuyển tiếp. Thay vào đó, mặt IRS có thể thích ứng hoạt động trong một hoặc nhiều khoảng thời gian khác nhau, tùy thuộc vào thiết kế đặc thù và yêu cầu của ứng dụng trong hệ thống truyền thông. Ngoài ra, IRS cung cấp những lợi ích nổi bật sau đây:

- Thiết kế IRS đơn giản và tiêu tốn năng lượng thấp.
- IRS hoạt động như một gương phản xạ đặc biệt nên hầu như ít bị tác động bởi nhiễu cộng.

- Hiệu quả phổ của IRS cao do không bị ảnh hưởng bởi các giao thức bán hai chiều và nhiễu tự vòng lặp. Ngoài ra, hệ số phản xạ của bề mặt vật liệu IRS có thể được thiết kế để tối ưu sự liên kết giữa các tín hiệu từ máy phát và IRS.
- ❖ Đối với công nghệ mạng 5G và mạng tương lai, IRS có một vai trò rất hết sức trọng do nhiều lợi ích của IRS đem lại như sau [75]:
 - Dễ triển khai: IRS là các thiết bị gần như bị động, được làm từ vật liệu điện từ nên IRS có thể triển khai trên nhiều cấu trúc như mặt tiền tòa nhà, tường trong nhà, bảng quảng cáo ven đường, cột đường cao tốc và cửa sổ xe hơi.
 - Nâng cao hiệu suất phổ: IRS có thể tái cấu trúc môi trường truyền dẫn không dây, giảm thiểu sự giảm công suất ở khoảng cách xa.
 - Thân thiện với môi trường: IRS có khả năng điều chỉnh tín hiệu đến qua việc kiểm soát sự biến đổi pha của từng yếu tố phản xạ, thay cho việc dùng bộ khuếch đại công suất.
 - Khả năng tương thích: IRS hỗ trợ truyền hai chiều và sử dụng toàn bộ băng thông do IRS có khả năng phản xạ các sóng điện từ. Ngoài ra, IRS cải thiện hiệu năng cho phần lớn các thành phần mạng không dây, bảo đảm tuân thủ các tiêu chuẩn quốc tế và tương thích với phần cứng của mạng không dây mới nhất.

2.4 Kỹ thuật đa truy nhập phi trực giao miền công suất (PD-NOMA)

NOMA được coi là kỹ thuật đa truy nhập tiên tiến trong mạng thế hệ mới 5G. Đặc biệt, một loại cơ bản quan trọng của NOMA là ghép kênh trong miền công suất (PD-NOMA) đang được sự chú ý gần đây. PD-NOMA cho phép phục vụ đồng thời nhiều người dùng trên cùng một kênh thời gian, tần số, và mã, với điều kiện là phân chia công suất biến đổi tùy theo chất lượng của kênh. Người dùng với trạng thái kênh tốt hơn sẽ nhận mức tiêu hao ít công suất hơn. Bản chất cốt lõi của PD-NOMA là kết hợp các tín hiệu một cách tuyến tính và truyền đồng thời tín hiệu đã kết hợp này đến đích dựa vào kỹ thuật mã hóa xếp chồng (SC). Tại phía nguồn nhận, PD-NOMA tận dụng sự chênh lệch công suất của các tín hiệu để phân biệt

các người dùng, thông qua kỹ thuật loại bỏ nhiễu một cách tuần tự (SIC) [76]. Hai kỹ thuật này được trình bày ngắn gọn như sau:

❖ Kỹ thuật SC

Kỹ thuật mã hóa xếp chồng SC cho phép truyền dữ liệu từ một nguồn đến nhiều người nhận đồng thời. Bằng cách mã hóa riêng biệt cho từng người dùng và xếp chồng các tín hiệu lên nhau, kỹ thuật này tối ưu hóa hiệu quả truyền. Trong trường hợp hai người dùng, nguồn sẽ sử dụng hai bộ mã hóa điểm-điểm, ký hiệu là $f_1: \{0,1\}^{[2^{kR_1}]} \rightarrow C^k$ và $f_2: \{0,1\}^{[2^{kR_2}]} \rightarrow C^k$. Mỗi bộ mã hóa này sẽ ánh xạ các bit đầu vào thành chuỗi tín hiệu đầu ra $x_1(n)$ và $x_2(n)$ với chiều dài khối là k . Ở đây, R_1 và R_2 lần lượt là tốc độ truyền của người dùng 1 và người dùng 2, và $[\cdot]$ là toán tử lấy phần nguyên. C là một thư viện mã. Cuối cùng, thiết bị tổng hợp sẽ cung cấp một chuỗi đầu ra:

$$x(n) = \sqrt{\alpha_1 P} x_1(n) + \sqrt{\alpha_2 P} x_2(n), \quad (2.30)$$

trong đó α_i hệ số phân bổ công suất được phân bổ cho người dùng $i = \{1, 2\}$, với điều kiện: $\sum_i \alpha_i = 1$. Trong quá trình truyền tín hiệu, tín hiệu của người dùng 1 và người dùng 2 sẽ được truyền đồng thời nhưng với công suất khác nhau. Điều này giúp bộ thu có thể giải mã từng tín hiệu riêng biệt nhờ vậy mà tín hiệu xếp chồng được phân tách một cách hiệu quả trong không gian tần số.

❖ Kỹ thuật SIC

Kỹ thuật SIC trong truyền thông không dây hoạt động dựa trên ý tưởng giải mã tuần tự để loại bỏ nhiễu giữa các tín hiệu chồng chéo từ nhiều người dùng. Mỗi bước của quy trình SIC được thực hiện như sau: các tín hiệu được sắp xếp từ mạnh đến yếu, bộ thu sẽ giải mã tín hiệu mạnh nhất trước và loại bỏ nó khỏi tín hiệu tổng hợp để giảm nhiễu cho các tín hiệu còn lại. Tiếp đó, bộ thu lần lượt giải mã các tín

hiệu yếu hơn theo thứ tự, đảm bảo độ chính xác cho từng tín hiệu. Về mặt toán học, quy trình SIC cho hai người dùng được mô tả cụ thể như sau:

1) Người dùng 1: giải mã tín hiệu

Bộ giải mã đơn $g1: C^k \rightarrow \{0,1\}^{[2^{kR_1}]}$ sẽ giải mã tín hiệu $x_1(n)$ bằng cách xem $x_2(n)$ là nhiễu.

2) Người dùng 2: giải mã tín hiệu tuần tự.

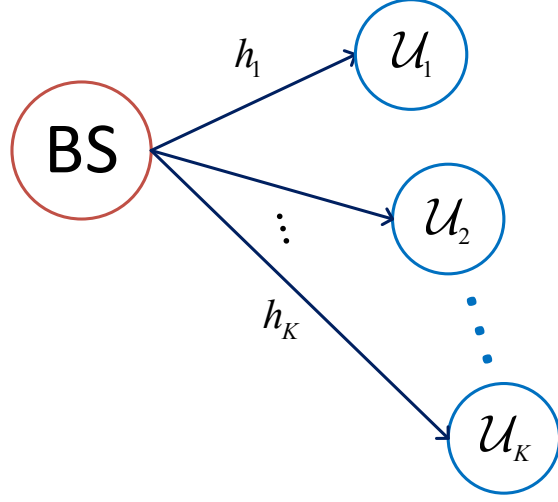
- Bước 1: Giải mã tín hiệu của người dùng 1, $x_1(n)$, bằng cách sử dụng bộ giải mã đơn $g1: C^k \rightarrow \{0,1\}^{[2^{kR_1}]}$.
- Bước 2: trừ phần nhiễu của tín hiệu người dùng 1, $\sqrt{\alpha_1 P} h_2 x_1(n)$, khỏi tín hiệu nhận được $Y_2(n)$, bằng phép tính: $Y'_2(n) = Y_2(n) - \sqrt{\alpha_1 P} h_2 x_1(n)$, trong đó h_2 hệ số kênh phức tại người dùng 2.
- Bước 3: giải mã tín hiệu $x_2(n)$ của người dùng 2 bằng cách áp dụng một bộ giải mã đơn khác $g2: C^k \rightarrow \{0,1\}^{[2^{kR_2}]}$ trên đầu vào là $Y'_2(n)$. Quy trình này giúp người dùng 2 loại bỏ tín hiệu của người dùng 1 và chỉ còn lại tín hiệu của mình, giảm thiểu nhiễu và cải thiện độ chính xác của tín hiệu nhận được.

Tiếp theo, mô hình truyền thông cơ bản sử dụng PD-NOMA được minh họa qua Hình 2.12. Nguồn BS muốn gửi cùng lúc dữ liệu x_i , $i \in K = \{1, 2, \dots, K\}$, đến K người dùng $\mathcal{U}_i$, tương ứng. Tất cả nguồn BS và các người dùng sử dụng an-ten đơn để truyền và nhận tín hiệu. Đầu tiên, nguồn BS dùng kỹ thuật SC để ghép các dữ liệu x_i này lại cùng nhau và sau đó gửi đến cho mọi người dùng $\mathcal{U}_i$. Tín hiệu kết hợp được biểu diễn như sau [77]:

$$x_{BS} = \sum_{i=1}^K \sqrt{\alpha_i P} x_i, \quad (2.31)$$

với các hệ số phân bổ công suất sẽ được đặt theo thứ tự như sau:

$0 < \alpha_1 < \alpha_2 < \dots < \alpha_K$ và $\sum_{i=1}^K \alpha_i = 1$; tổng công suất của nguồn BS là $P = \sum_{i=1}^K P_i$.



Hình 2.12: Mô hình truyền thông sử dụng PD-NOMA với K người dùng.

Tiếp theo, các người dùng $\mathcal{U}_i$ dựa vào kỹ thuật SIC để lần lượt khử các thành phần tín hiệu khác để thu tín hiệu mong muốn. Cụ thể, mỗi người dùng $\mathcal{U}_i$ sẽ thực hiện quá trình giải mã xử lý tín hiệu từ người dùng yếu, tức là $\mathcal{U}_i$ sẽ ưu tiên giải mã tín hiệu của người dùng $\mathcal{U}_K$ trước với $K > i$. Sau đó, $\mathcal{U}_i$ sẽ loại bỏ tín hiệu của $\mathcal{U}_K$ bằng cách coi tín hiệu này như nhiễu nội và tiếp tục giải mã tín hiệu của người dùng $\mathcal{U}_{K-1}$ còn lại. Quá trình này được lặp đi lặp lại cho đến khi $\mathcal{U}_i$ giải mã tín hiệu của chính mình thành công. Tín hiệu mà người dùng $\mathcal{U}_i$ nhận được có thể được mô tả sau đây:

$$y_{\mathcal{U}_i} = h_i x_{BS} + n_i, \quad (2.32)$$

với h_i là hệ số kênh truyền fading của đường truyền từ BS đến $\mathcal{U}_i$; n_i là nhiễu trắng Gauss tại người dùng $\mathcal{U}_i$, với trung bình bằng 0 và phương sai σ_i^2 , tương ứng. Nếu SC tại BS và SIC tại $\mathcal{U}_i$ là hoàn hảo, tỉ số tín hiệu cộng nhiễu (SINR) tại người dùng $\mathcal{U}_i$ như sau:

$$\gamma_{u_i} = \frac{\alpha_i P |h_i|^2}{P |h_i|^2 \sum_{j=i+1}^K \alpha_j + \sigma_i^2}, \quad (2.33)$$

Tại $\mathcal{U}_K$ sau khi loại bỏ hết các tín hiệu khác trước nó thành công, SINR được tính như sau:

$$\gamma_{u_K} = \frac{\alpha_K P |h_K|^2}{\sigma_K^2}. \quad (2.34)$$

Hiện nay, so với phương pháp đa truy nhập trực giao (OMA) trước đây, NOMA vẫn được nghiên cứu tích cực như một giải pháp hứa hẹn cho 5G và các công nghệ mạng tương lai bởi nhiều ưu thế vượt trội như sau [77]:

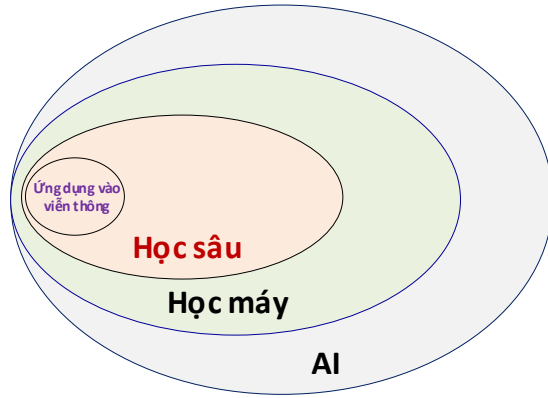
- Hiệu quả phổ và thông lượng được tăng cường: trong NOMA, tài nguyên thời gian, tần số được chia sẻ không trực giao giữa người dùng cả trong miền công suất và trong miền mã. Do đó, NOMA có khả năng đạt được dung lượng tổng tối đa và hỗ trợ công bằng người dùng hơn trong đường lên của các kênh AWGN. Ngoài ra, giới hạn dung lượng của NOMA cao hơn so với OMA trong đường xuống của các kênh AWGN.
- Đa kết nối: Phân bổ tài nguyên không trực giao trong NOMA cho thấy rằng số người dùng/thiết bị có thể hỗ trợ không bị giới hạn bởi số lượng tài nguyên trực giao có sẵn. Do đó, NOMA có tiềm năng hỗ trợ kết nối khối lượng lớn.
- Độ trễ truyền tín hiệu thấp: Trong NOMA, việc truyền đường lên không cần cấp quyền do đó giảm đáng kể cả độ trễ truyền và chi phí tín hiệu truyền.
- Phản hồi kênh giảm: Yêu cầu về phản hồi kênh sẽ được giảm nhẹ trong NOMA theo miền công suất, vì thông tin phản hồi CSI chỉ được sử dụng cho phân bổ công suất. Do đó, không cần thiết phải có thông tin CSI chính xác ngay lập tức. Dù người dùng là cố định hay di động, việc có một phản hồi kênh với độ chính xác hạn chế và độ trễ tối đa nhất định sẽ không gây ra ảnh hưởng lớn đối với hiệu năng hệ thống có thể đạt được, miễn là kênh không thay đổi nhanh chóng.

2.5 Kỹ thuật học sâu

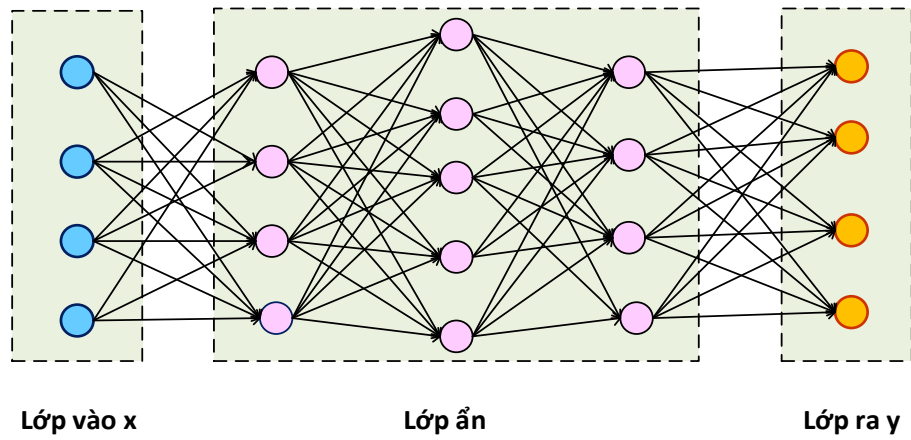
Học sâu DL là một nhánh con của ML thuộc AI được mô tả qua Hình 2.13. Cụ thể, AI là một mô hình tính toán cung cấp khả năng thông minh cho máy, nhằm giúp chúng học cách làm việc, phản ứng và học hỏi như con người. Nhiều kỹ thuật của AI bao gồm học máy, hệ thống chuyên gia và các thuật toán tiến hóa. Trong đó, học máy cho phép quá trình nhân tạo hấp thụ thông tin từ dữ liệu và đưa ra quyết định mà không cần được lập trình một cách rõ ràng. Tập con quan trọng của ML, đặc biệt là DL, cũng có khả năng thực hiện các công việc như dự đoán, phân loại hoặc tiến hành ra quyết định dựa vào dữ liệu mà không yêu cầu lập trình phức tạp. Một số thuật toán thông dụng cho DL như hồi quy tuyến tính, bộ phân loại K-nearest và Q-learning. Mô hình học sâu phổ biến là các mạng nơ ron nhiều lớp (LMPs) được minh họa như Hình 2.14. Mạng nơ ron nhiều lớp MLPs là thiết kế mạng nơ ron nhân tạo ban đầu, gồm ít nhất ba lớp hoạt động như lớp đầu vào, các lớp ẩn và lớp đầu ra [175].

Trong mạng MLPs, dữ liệu tự nhiên được đưa vào hệ thống học ở dạng dữ liệu thô. Sau đó, hệ thống học sâu DL tự động trích xuất các biểu diễn phù hợp cho mục đích phân loại hoặc phát hiện. Bắt đầu từ dữ liệu tự nhiên, mỗi lớp rút trích các đặc trưng độc lập từ dữ liệu đầu vào, từng bước tăng cường các đặc trưng liên quan đến quyết định và giảm thiểu các đặc trưng không liên quan. Mỗi lớp kết nối với các lớp láng giềng với các trọng số khác nhau được gắn kết vào kết nối. Để tìm giá trị cho các trọng số, một số lượng lớn mẫu được đưa vào hệ thống cho mục đích huấn luyện, có thể là học có giám sát (SL) hoặc không giám sát (UL). Trong SL, một vector độ dốc được tính toán cho mỗi trọng số, chỉ ra lượng thay đổi lỗi theo biến thiên của trọng số đó. Dựa vào vec tơ độ dốc, trọng số được điều chỉnh để giảm thiểu lỗi.

Nguyên lý cơ bản của học sâu gồm hai quá trình quan trọng: quá trình lan truyền tiến và phản hồi lỗi ngược [78]. Trong quá trình huấn luyện luôn cần cả hai quá trình này.



Hình 2.13: Sơ đồ Venn mô tả mối quan hệ giữa DL, ML và AI.



Hình 2.14: Cấu trúc của MLPs với 3 lớp ẩn.

- **Lan truyền tiến** là quá trình mà qua đó mạng nơ ron biến đổi dữ liệu đầu vào từ dạng thô sang biểu diễn dữ liệu trừu tượng. Điều này cải thiện đáng kể khả năng của mạng trong việc ra quyết định và dự đoán một cách chính xác và hiệu quả. Ví dụ, trong Hình 2.15 . Gọi n_{ij} là nơ ron (nút) thứ j trong lớp i , đầu ra của nút n_{ij} xác định được thông qua hai giai đoạn. Ban đầu, tính tổng các trọng số của tất cả đầu vào của nó, được ký hiệu là z_{ij} . Sau đó, z_{ij} được gửi đến một hàm phi tuyến $f(\cdot)$ để thu được đầu ra y_{ij} của nút n_{ij} như sau:

$$y_{ij} = f(z_{ij}), \quad (2.35)$$

$$z_{ij} = \sum_{k=1}^{L_{i-1}} w_{kj}^i, \quad (2.36)$$

với w_{kj} là trọng số từ nút $n_{i-1,k}$ đến nút n_{ij} và L_{i-1} là số nút ở lớp $i-1$; $f(.)$ là hàm kích hoạt của MLPs và có vai trò cốt lõi trong việc mang lại đặc tính phi tuyến cho mô hình, và giúp nâng cao khả năng học cho mô hình và xử lý được các dữ liệu phức tạp. Những hàm kích hoạt phổ biến được sử dụng rộng rãi như Sigmoid, ReLU, Tanh và Softmax [79].

- **Phản hồi lỗi ngược:** Đây là phương pháp tối ưu hóa trọng số để giảm thiểu sai số giữa đầu ra dự đoán và giá trị thật. Ban đầu, trọng số của các kết nối được chọn ngẫu nhiên. Sau đó, sai số giữa giá trị dự đoán và giá trị thực được tính toán tại lớp đầu ra, ví dụ như tại nút n_{N_j} đạo hàm của hàm mất mát E (minh họa qua Hình 2.16) được tính như sau:

$$\frac{\delta E}{\delta z_{N_j}} = \frac{\delta E}{\delta y_{N_j}} \frac{\delta y_{N_j}}{\delta z_{N_j}}, \quad (2.37)$$

với $\frac{\delta E}{\delta y_{N_j}} = y_{N_j} - t_{N_j}$ và $j = 1, 2, \dots, L_N$.

Đạo hàm này sau đó được lan truyền ngược qua các lớp để điều chỉnh trọng số của các kết nối ở lớp thấp hơn. Ví dụ, đạo hàm tại các nơ ron lớp thấp hơn n_{ij} được tính như sau:

$$\frac{\delta E}{\delta z_{ij}} = \frac{\delta E}{\delta y_{ij}} \frac{\delta y_{ij}}{\delta z_{ij}}, \quad (2.38)$$

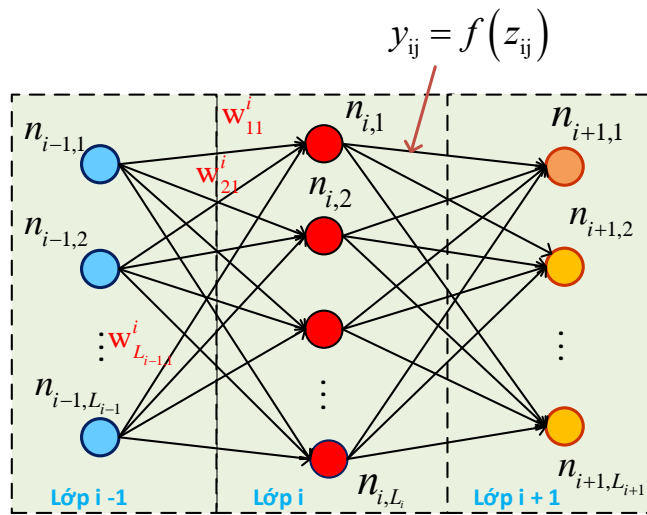
với $\frac{\delta E}{\delta y_{ij}} = \sum_{k=1}^{L_{i+1}} w_{jk}^{i+1} \frac{\delta E}{\delta z_{i+1,k}}$, $i = 1, 2, \dots, N-1$ và $j = 1, 2, \dots, L_i$.

Hiện nay, học sâu có tầm quan trọng rất lớn trong nhiều ngành công nghệ, đặc biệt trong viễn thông nhằm nâng cao hiệu suất mạng theo thời gian. Nhiều ưu điểm nổi trội mà học sâu mang lại như sau:

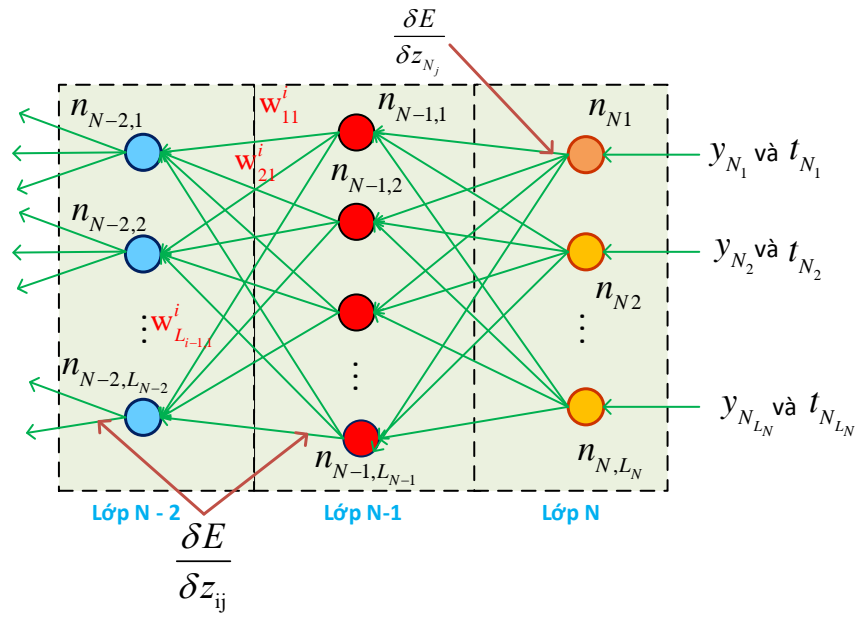
- Học sâu có thể rút trích các đặc trưng của dữ liệu mức độ cao, có cấu trúc và mối tương quan phức tạp. Quá trình học không cần được thiết kế bởi con người.

- Học sâu có thể xử lý một lượng dữ liệu lớn giúp cho việc phân tích và dự đoán hiệu quả hơn.
- Học sâu cung cấp nhiều phương pháp cho phép khai thác dữ liệu không được gán nhãn để học các mô hình mẫu hữu ích một cách không giám sát.
- Học sâu có thể chia sẻ biểu diễn: mạng nơ ron sâu có thể được chia sẻ trên các nhiệm vụ khác nhau, giúp giảm yêu cầu tính toán và bộ nhớ của hệ thống di động khi thực hiện các ứng dụng học nhiều nhiệm vụ.

Học sâu có thể xử lý dữ liệu hình học hiệu quả, bao gồm thông tin về vị trí của người dùng di động cũng như hình ảnh và video từ thiết bị di động.



Hình 2.15: Quá trình lan truyền tiến của DL.



Hình 2.16: Quá trình phản hồi lỗi ngược của DL.

Chương 3. TỐI ƯU HIỆU NĂNG HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG CHUYỂN TIẾP HAI CHẶNG

3.1 Giới thiệu

Trọng tâm chính của Chương 3 là thực hiện mục tiêu đầu tiên của Luận án, tập trung vào việc khai thác hiệu quả độ phân tập không gian mạng và tối ưu hiệu năng mạng, qua đó góp phần hiện thực hệ thống uRLLCs. Gần đây, các nghiên cứu về SPC trong mạng chuyển tiếp đã được quan tâm đáng kể. Cụ thể, công trình [80] nghiên cứu mô hình chuyển tiếp AF đơn giản với ba nút, bao gồm hai nút nguồn và một nút chuyển tiếp, với hai phương án truyền qua hai khe thời gian (2TS) và ba khe thời gian (3TS). Kết quả cho thấy phương án 3TS phù hợp hơn với hệ thống có tốc độ truyền thông tin thấp và yêu cầu độ trễ ít nghiêm ngặt hơn so với 2TS. Tiếp theo, nghiên cứu SPC được mở rộng cho mạng đa chặng chuyển tiếp DF qua công trình [41, 81]. Các tác giả trong những nghiên cứu này đã chứng minh rằng mạng chuyển tiếp hai chặng đạt hiệu năng tối ưu. Tiếp tục mở rộng nghiên cứu, các công trình SPC trong mạng chuyển tiếp hai chặng đã khai thác độ phân tập không gian mạng, tiêu biểu như các công trình [19], [24] và [82]. Cụ thể, công trình [19] đề xuất giao thức PRS và phân tích hiệu năng hệ thống thông qua BLER phụ thuộc vào SNR, chiều dài khối, và bậc phân tập được xác định bởi số lượng nút chuyển tiếp trong mạng. Công trình [24] giới thiệu các giao thức MRC và SC áp dụng cho nguồn phát và nguồn nhận tín hiệu, đồng thời chứng minh rằng giao thức MRC cải thiện hiệu năng đáng kể so với SC. Công trình [82] kết hợp và mở rộng với hai giao thức PRS và MRC, với phân tích cho thấy sự cải thiện về độ phân tập phụ thuộc vào số nút chuyển tiếp N và số an-ten M tại nguồn nhận.

Tuy nhiên, hầu hết các nghiên cứu hiện tại chủ yếu tập trung vào việc phân tích, đánh giá độ tin cậy hệ thống dựa trên tốc độ truyền và BLER, chưa có nghiên cứu nào đánh giá toàn diện hệ thống cả độ tin cậy và độ trễ trong các tình huống phức tạp.

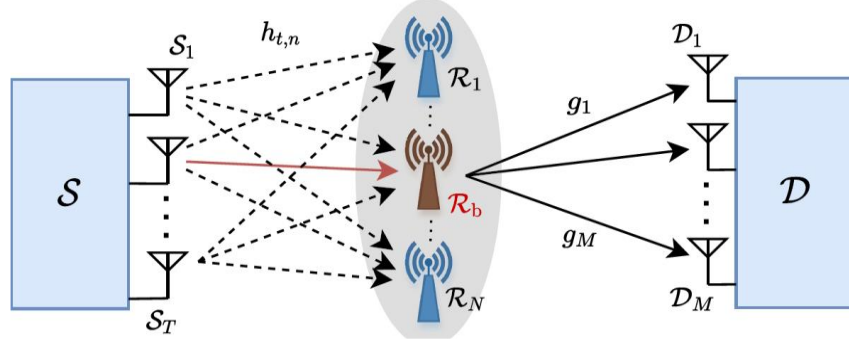
Dựa trên cơ sở của các nghiên cứu trên, Chương 3 của Luận án này đề xuất một mô hình khai thác độ phân tập không gian bằng cách kết hợp ba giao thức TAS, PRS và MRC. Mô hình này xác định độ phân tập đạt được là $\min(NT, M)$, với T là số an-ten phát, từ đó đóng vai trò quan trọng trong thiết kế hệ thống mạng thực tế. Mô hình này không chỉ giúp giảm thiểu chi phí triển khai mà còn nâng cao đáng kể độ tin cậy của hệ thống. Hơn nữa, để đáp ứng các yêu cầu khắt khe của uRLLC, Luận án còn đề xuất một bài toán tối ưu hóa công suất phát và vị trí các nút chuyển tiếp, nhằm cải thiện hiệu năng hệ thống.

Chương 3 được tổ chức thành các mục chính sau đây: Mục 3.1 giới thiệu của chương. Mục 3.2 trình bày mô hình nghiên cứu mới, trong đó áp dụng các giao thức như TAS, PRS và MRC cho hệ thống. Mục 3.3 phân tích và tính toán các tham số hiệu năng của hệ thống như BLER, độ trễ và thông lượng qua kênh truyền fading Rayleigh. Mục 3.4 chú trọng vào việc nâng cao hiệu năng của hệ thống bằng cách xây dựng bài toán tối ưu công suất nguồn phát và vị trí của các nút chuyển tiếp, đồng thời đề xuất phương pháp giải bài toán tối ưu dưới hai điều kiện ràng buộc: công suất phân bổ nguồn cố định và chuẩn hóa khoảng cách giữa nguồn phát và nguồn thu. Mục 3.5 trình bày các kết luận từ phân tích và mô phỏng Monte Carlo, đồng thời xác nhận rằng hiệu năng trong tình huống được tối ưu đem lại cải thiện nhiều so với tình huống không được tối ưu.

3.2 Mô hình nghiên cứu

Hình 3.1 minh họa một mô hình SPC trong mạng DF hai chặng, bao gồm một nút nguồn ($\mathcal{S}$) có $\mathcal{S}_t$ an-ten với $t = 1, 2, \dots, T$, N nút chuyển tiếp ($\mathcal{R}_n$) với $n = 1, \dots, N$ sử dụng một an-ten và nút đích ($\mathcal{D}$) có M an-ten $\mathcal{D}_m$ với $m = 1, \dots, M$. Chúng ta giả định rằng không có kết nối trực tiếp giữa nút $\mathcal{S}$ và nút $\mathcal{D}$. Hệ thống hoạt động dựa trên cơ chế phân chia thời gian cho đa truy nhập với hai khe thời gian liên tiếp qua các nút $\mathcal{R}_n$. Trong khe thời gian thứ nhất, nút $\mathcal{S}$ sẽ truyền tín hiệu đến các nút $\mathcal{R}_n$, trong đó hệ thống sử dụng đồng thời các kỹ thuật TAS và

PRS. Theo nguyên tắc TAS, hệ thống chỉ chọn một an-ten phát tối ưu của nút nguồn để thực hiện quá trình truyền, tức là một an-ten có giá trị SNR cao nhất được chọn.



Hình 3.1: Mô hình mạng DF hai chặng với các kỹ thuật TAS/PRS/MRC trong SPC.

Thông qua việc áp dụng kỹ thuật PRS, nút $\mathcal{S}$ sẽ lựa chọn ra một nút từ tập N nút của $\mathcal{R}_n$ dựa trên giá trị SNR cao nhất. Gọi $\mathcal{R}_b$ là nút chuyển tiếp được chọn với mục đích truyền và xử lý tín hiệu từ nút $\mathcal{S}$ đến nút $\mathcal{D}$ [83] với giá trị b biểu thị như sau:

$$b = \arg \max_{\substack{t=1, \dots, T \\ n=1, \dots, N}} \gamma_{\mathcal{S}\mathcal{R}_{t,n}}, \quad (3.1)$$

với $\gamma_{\mathcal{S}\mathcal{R}_{t,n}}$ là SNR của liên kết từ $\mathcal{S}_t \rightarrow \mathcal{R}_n$. Ở đây, kênh truyền giữa nút $\mathcal{S}$ đến nút $\mathcal{D}$ được giả định là kênh fading Rayleigh bán tĩnh [84]. Do đó, $\gamma_{\mathcal{S}\mathcal{R}_{t,n}}$ tuân theo phân phối mũ và giá trị trung bình như sau:

$$\bar{\gamma}_{\mathcal{S}\mathcal{R}_{t,n}} = \frac{P_S}{\sigma^2} \mathbb{E} \left\{ |h_{t,n}|^2 \right\} = \frac{P_S \Omega_{\mathcal{S}\mathcal{R}_{t,n}}}{\sigma^2}, \quad (3.2)$$

với P_S là công suất của nguồn $\mathcal{S}$; $h_{t,n}$ và $\Omega_{\mathcal{S}\mathcal{R}_n}$ là các hệ số và độ lợi trung bình cho kênh truyền từ $\mathcal{S}_t \rightarrow \mathcal{R}_n$, tương ứng. σ^2 là công suất nhiễu Gauss AWGN tại $\mathcal{R}_n$. Hàm CDF của SNR $\gamma_{\mathcal{S}\mathcal{R}_b}$ đầu ra trong chặng thứ nhất được cho như sau:

$$F_{\gamma_{SR_b}}(\gamma) = \prod_{t=1}^T \prod_{n=1}^N \Pr(\gamma_{SR_{t,n}} \leq \gamma) = \prod_{t=1}^T \prod_{n=1}^N F_{\gamma_{SR_{t,n}}}(\gamma). \quad (3.3)$$

Hơn nữa, hệ thống cũng giả định rằng các kênh từ $\mathcal{S}_t \rightarrow \mathcal{R}_n$ và $\mathcal{R}_b \rightarrow \mathcal{D}_m$ là các kênh fading Rayleigh độc lập và đồng dạng nhau (i.i.d.). Do đó, SNR trung bình của từng nhánh là như nhau, tức $\bar{\gamma}_{SR_{t,n}} = \bar{\gamma}_{SR}$.

Bằng cách sử dụng phép khai triển nhị thức theo [85], hàm CDF của γ_{SR_b} có thể tính được bên dưới đây:

$$\begin{aligned} F_{\gamma_{SR_b}}(\gamma) &= \left[1 - \exp\left(-\frac{\gamma}{\bar{\gamma}_{SR}}\right) \right]^{NT} \\ &= \sum_{n=1}^N (-1)^{n-1} \binom{NT}{n} \left[1 - \exp\left(-\frac{n\gamma}{\bar{\gamma}_{SR}}\right) \right]. \end{aligned} \quad (3.4)$$

Tiếp theo, trong khe thời gian thứ hai, $\mathcal{R}_b$ được chọn sẽ thu nhận tín hiệu từ nút $\mathcal{S}$, tiếp theo đó, $\mathcal{R}_b$ sẽ tiến hành xử lý và chuyển tiếp tín hiệu này đến nút $\mathcal{D}$ bằng phương pháp DF. Để hiệu quả và độ tin cậy cho việc nhận tín hiệu, nút $\mathcal{D}$ triển khai giao thức MRC. Do đó, tỉ lệ SNR đầu ra tại nút $\mathcal{D}$ có thể được thu được sau đây:

$$\gamma_{\mathcal{R}_b\mathcal{D}} = \sum_{m=1}^M \gamma_{\mathcal{R}_b\mathcal{D}_m}, \quad (3.5)$$

với $\gamma_{\mathcal{R}_b\mathcal{D}_m}$ là SNR nhận tại $\mathcal{D}_m$ và giá trị trung bình SNR trong mỗi nhánh là

$$\bar{\gamma}_{\mathcal{R}_b\mathcal{D}_m} = \bar{\gamma}_{\mathcal{R}_b\mathcal{D}} = \frac{P_{\mathcal{R}_b}}{\sigma^2} \Omega_{\mathcal{R}_b\mathcal{D}} \text{ với } P_{\mathcal{R}_b} \text{ là công suất của nút } \mathcal{R}_b \text{ và } \Omega_{\mathcal{R}_b\mathcal{D}} \text{ là độ lợi trung}$$

bình của kênh truyền từ $\mathcal{R}_b \rightarrow \mathcal{D}$, tương ứng. Hơn nữa, độ lợi kênh truyền từ

$\mathcal{R}_b \rightarrow \mathcal{D}_m$, hay $|g_m|^2$ với $(m=1, \dots, M)$, có phân bố mũ. Do đó, tỉ lệ SNR $\gamma_{\mathcal{R}_b\mathcal{D}_m}$

theo phân phối Chi bình phương với kỳ vọng là $M\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}$ và phương sai là $2M\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}$

[86]. Trong trường hợp này, hàm CDF của $\gamma_{\mathcal{R}_b\mathcal{D}}$ nhận được như sau:

$$F_{\gamma_{\mathcal{R}_b\mathcal{D}}}(\gamma) = 1 - \exp\left(-\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}}\right) \sum_{m=1}^M \frac{\left(\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}}\right)^{m-1}}{(m-1)!}. \quad (3.6)$$

3.3 Phân tích hiệu năng hệ thống

Trong phần này, chúng ta sẽ xem xét hiệu năng của hệ thống qua các chỉ số như BLER, thông lượng và độ trễ. Để thuận tiện hơn trong phân tích hiệu năng, chúng ta tiến hành phân tích BLER dạng xấp xỉ ở cả mức SNR trung bình và cao. Xấp xỉ BLER này chính là cận trên tối đa của BLER khi giá trị trung bình của SNR tăng lên đến vô cùng. Do đó, BLER của hệ thống sẽ luôn nằm dưới hoặc bằng giới hạn trên, bất kể giá trị SNR trung bình tăng lên đến mức nào.

3.3.1 Tỷ lệ lỗi khối trung bình

Giả sử hệ thống có nút nguồn $\mathcal{S}$ truyền β bit thông tin đến nút đích $\mathcal{D}$ trong hai khoảng thời gian với tổng độ dài khối tin k . Khối tin này được chia đều cho mỗi khoảng thời gian, và tốc độ mã hóa thông tin cho mỗi khe được duy trì ở mức $\mathcal{R} = \frac{2\beta}{k}$.

Thay (3.4) và (3.6) vào (2.9), BLER trung bình ở hai chặng thu được lần lượt sau đây:

$$\bar{\varepsilon}_{\mathcal{SR}_b} \approx 1 + \nu \sqrt{k} \sum_{n=1}^{NT} (-1)^{n-1} \binom{NT}{n} \left[\frac{\bar{\gamma}_{\mathcal{SR}}}{n} \left(\exp\left(-\frac{n\rho_H}{\bar{\gamma}_{\mathcal{SR}}}\right) - \exp\left(-\frac{n\rho_L}{\bar{\gamma}_{\mathcal{SR}}}\right) \right) \right]. \quad (3.7)$$

$$\bar{\varepsilon}_{\mathcal{R}_b\mathcal{D}} \approx 1 - \nu \sqrt{k} \sum_{m=1}^M \frac{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}}{(m-1)!} \left\{ \Upsilon\left(m, \frac{\rho_H}{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}}\right) - \Upsilon\left(m, \frac{\rho_L}{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}}\right) \right\}. \quad (3.8)$$

Từ (3.7) và (3.8), BLER toàn trình (e2e BLER) được xét trong hai tình huống giải mã ở các nút chuyển tiếp sau đây:

- Tình huống thứ nhất là nút $\mathcal{R}_b$ áp dụng kỹ thuật SDF thì e2e BLER trung bình thu được sau đây:

$$\bar{\mathcal{E}}_{e2e}^{\text{SDF}} = \bar{\mathcal{E}}_{\mathcal{SR}_b} + (1 - \bar{\mathcal{E}}_{\mathcal{SR}_b}) \bar{\mathcal{E}}_{\mathcal{R}_b \mathcal{D}}. \quad (3.9)$$

- Tình huống thứ hai là nút $\mathcal{R}_b$ áp dụng kỹ thuật FDF thì e2e BLER trung bình sẽ được tính bên dưới như sau:

$$\bar{\mathcal{E}}_{e2e}^{\text{FDF}} = \bar{\mathcal{E}}_{\mathcal{SR}_b} (1 - \bar{\mathcal{E}}_{\mathcal{R}_b \mathcal{D}}) + (1 - \bar{\mathcal{E}}_{\mathcal{SR}_b}) \bar{\mathcal{E}}_{\mathcal{R}_b \mathcal{D}}. \quad (3.10)$$

Tiếp theo, nhằm đánh giá một cách đầy đủ cho hiệu năng của hệ thống, các chỉ số như độ trễ và thông lượng toàn trình của hệ thống truyền SPC được xem xét như sau:

a. Độ trễ toàn trình

Độ trễ toàn trình của hệ thống bao gồm nhiều yếu tố như độ trễ xử lý tại nguồn phát và nguồn thu, độ trễ hàng đợi, độ trễ căn chỉnh gói tin và độ trễ truyền gói tin. Trong đó, độ trễ truyền gói tin, được hiểu là khoảng thời gian cần thiết để gói tin di chuyển từ nguồn đến đích, đặc trưng chủ yếu cho độ trễ toàn trình của hệ thống. Khoảng thời gian truyền tin này phụ thuộc vào độ dài khối tin và thời gian di chuyển của tín hiệu trong không gian [87].

Gọi $L(k)$ là độ trễ giải mã một gói tin với độ dài khối k . Dựa vào [25, CT. (6)], độ trễ toàn trình của hệ thống (được đo bằng số CUs) được tính sau đây:

$$\begin{aligned} \tau &= \underbrace{\left[k + L(k) (1 - \bar{\mathcal{E}}_{\mathcal{SR}_b}) \right]}_{\text{chàng dàu}} + \underbrace{\left[k + L(k) \right]}_{\text{chàng thu hai}} \\ &= k + L(k) (2 - \bar{\mathcal{E}}_{\mathcal{SR}_b}), \end{aligned} \quad (3.11)$$

với $L(k) = \alpha k$ và α là hệ số trễ giải mã [88, bổ đề 6].

Độ trễ tổng cộng trong (3.11) được xác định từ thời điểm nút $\mathcal{R}_b$ giải mã tín hiệu thành công cho đến khi thông báo được gửi đến nút $\mathcal{D}$, bất kể tin nhắn này được nút $\mathcal{D}$ giải mã thành công ở nút đích hay không.

b. Thông lượng toàn trình

Thông lượng toàn trình được xác định là tỉ lệ giữa số lượng bit thông tin nhận được thành công tại nút đích và độ trễ truyền gói toàn trình. Do đó, thông lượng toàn trình cho hệ thống (đơn vị bits/ CUs) được suy ra như sau:

$$\delta = \frac{\beta(1 - \bar{\varepsilon}_{e2e})}{\tau}. \quad (3.12)$$

3.3.2 Tỉ lệ lỗi khối dạng xấp xỉ

Trong phần này, chúng ta xác định giới hạn trên của BLER trung bình khi SNR trung bình tăng lên đến vô cùng thông qua việc áp dụng hai phương pháp, bao gồm xấp xỉ vô cùng và tích phân Riemann như sau:

a. *Cách xấp xỉ thứ nhất:* Áp dụng xấp xỉ vô cùng cho hàm $1 - \exp(-x) \stackrel{x \rightarrow 0}{\sim} x$ vào (3.4), dạng xấp xỉ của BLER trong chặng đầu tiên được tính như sau:

$$\tilde{\varepsilon}_{SR_b} \approx v\sqrt{k} \int_{\rho_L}^{\rho_H} \left(\frac{\gamma}{\bar{\gamma}_{SR}} \right)^{NT} d\gamma = \frac{v\sqrt{k}}{NT+1} \left(\frac{\rho_H^{NT+1} - \rho_L^{NT+1}}{\bar{\gamma}_{SR}^{NT}} \right). \quad (3.13)$$

Ở chặng thứ hai, dựa vào [85, CT. (8.352.6)], hàm CDF trong (3.6) có thể rút gọn sau đây:

$$F_{\gamma_{R_bD}}(\gamma) = \frac{\Upsilon\left(M, \frac{\gamma}{\bar{\gamma}_{R_bD}}\right)}{(M-1)!}. \quad (3.14)$$

Áp dụng [85, CT. (8354.1)] cho tiệm cận của hàm gamma trong (3.14) khi $\frac{\gamma}{\bar{\gamma}_{R_bD}} \rightarrow 0$ được cho bên dưới như sau:

$$\begin{aligned}
Y\left(M, \frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right) &= \left(\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right)^M \left[1 - \exp\left(-\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right) \sum_{n=1}^{\infty} \frac{(-1)^n}{n!(M+n)} \left(\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right)^n \right] \\
&\stackrel{\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}} \rightarrow 0}{=} \frac{1}{M} \left(\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right)^M.
\end{aligned} \tag{3.15}$$

Thay (3.15) và (3.14) vào (2.9), dạng xấp xỉ của BLER trung bình trong chặng thứ hai được tính sau đây:

$$\tilde{\varepsilon}_{\mathcal{R}_b \mathcal{D}} \approx v\sqrt{k} \int_{\rho_L}^{\rho_H} \frac{1}{M!} \left(\frac{\gamma}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}}\right)^M d\gamma = \frac{v\sqrt{k}}{M!(M+1)} \left(\frac{\rho_H^{M+1} - \rho_L^{M+1}}{\bar{\gamma}_{\mathcal{R}_b \mathcal{D}}^M} \right). \tag{3.16}$$

Mặt khác, chúng ta thấy được rằng BLER trung bình trong mỗi chặng $\bar{\varepsilon}_{\mathcal{SR}_b}$ và $\bar{\varepsilon}_{\mathcal{R}_b \mathcal{D}}$ trở nên rất nhỏ khi công suất phát tăng lên, điều này được biểu thị qua $\bar{\varepsilon}_{\mathcal{SR}_b} \bar{\varepsilon}_{\mathcal{R}_b \mathcal{D}} \approx 0$. Do đó, dạng gần đúng của tỉ lệ lỗi khối BLER từ (3.13) và (3.16) có thể viết lại như sau:

$$\tilde{\varepsilon}_{e2e}^{SDF} \approx \tilde{\varepsilon}_{e2e}^{FDF} \approx \tilde{\varepsilon}_{\mathcal{SR}_b} + \tilde{\varepsilon}_{\mathcal{R}_b \mathcal{D}}. \tag{3.17}$$

b. Cách xấp xỉ thứ hai: Áp dụng tích phân Riemann.

Từ (2.7) và (2.8), giá trị của $\rho_H - \rho_L = 2\pi\sqrt{(2^{2\mathcal{R}} - 1)/k}$ là nhỏ bởi vì $\mathcal{R}$ càng nhỏ khi $k \geq 100$ CUs [28]. Trong trường hợp này, phép gần đúng tích phân Riemann được đề xuất sử dụng với $\int_c^d f(x)dx = (c-d)f\left(\frac{c+d}{2}\right)$. Bằng cách áp dụng xấp xỉ Riemann bậc nhất, các tỉ lệ lỗi khối BLERs ở hai chặng được xấp xỉ như sau:

$$\tilde{\varepsilon}_{\mathcal{SR}_b}^{(2)} \approx v\sqrt{k} \left(\frac{\theta}{\bar{\gamma}_{\mathcal{SR}}}\right)^{NT}. \tag{3.18}$$

$$\tilde{\varepsilon}_{\mathcal{R}_b\mathcal{D}}^{(2)} \approx \frac{1}{M!} \left(\frac{\theta}{\bar{\gamma}_{\mathcal{R}_b\mathcal{D}}} \right)^M. \quad (3.19)$$

Tương tự (3.17), BLER toàn trình dạng xấp xỉ được suy ra như sau:

$$\tilde{\varepsilon}_{e2e}^{SDF} \approx \tilde{\varepsilon}_{e2e}^{FDF} \approx \tilde{\varepsilon}_{SR_b}^{(2)} + \tilde{\varepsilon}_{\mathcal{R}_b\mathcal{D}}^{(2)}. \quad (3.20)$$

Từ công thức (2.9), BLER trung bình ở mỗi chặng $\bar{\varepsilon}_{\mathcal{X}}$ chỉ phụ thuộc vào hàm $F_{\gamma_{\mathcal{X}}}(\gamma)$. Khi $F_{\gamma_{\mathcal{X}}}(\gamma)$ trở nên quá phức tạp, BLER xấp xỉ dạng tường minh khó có thể thu được dựa vào phương pháp xấp xỉ đầu tiên. Do đó, áp dụng xấp xỉ tích phân Riemann làm đơn giản quá trình tính toán. Trong phần này, việc sử dụng xấp xỉ thứ hai không chỉ nhằm mục đích giảm thiểu độ phức tạp mà còn để xác minh độ chính xác của phương pháp xấp xỉ đầu tiên.

c. Bậc phân tập của hệ thống

Bậc phân tập không gian của hệ thống có thể thu được như sau [89]:

$$\begin{aligned} \mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\tilde{\varepsilon}_{e2e}^{(1)})}{\log(\bar{\gamma})} = -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\tilde{\varepsilon}_{e2e}^{(2)})}{\log(\bar{\gamma})} \\ &= \min(NT, M). \end{aligned} \quad (3.21)$$

(Công thức (3.21) được chứng minh trong Phụ lục A).

3.4 Tối ưu hiệu năng hệ thống

Vấn đề tối ưu công suất của nguồn phát và vị trí nút chuyển tiếp nhằm giảm BLER toàn trình đến mức tối đa, đáp ứng yêu cầu của hệ thống uRLLCs được phân tích chi tiết bên dưới như sau:

3.4.1 Tối ưu công suất nguồn phát

Chúng ta giả định rằng hệ thống được cấp cho một công suất nguồn phát cố định. Một thách thức xuất phát là làm thế nào để tối ưu phân phối công suất giữa nút nguồn S và nút $\mathcal{R}_b$ trong tình huống giả định cố định vị trí của nút $\mathcal{R}_b$. Vấn đề tối ưu công suất được đặt ra như sau:

$$\min_{\mathbf{P}} \bar{\varepsilon}_{e2e}(\mathbf{P}) \text{ s.t. } P_S + P_{\mathcal{R}_b} \leq P_0, \quad (3.22)$$

với P_0 là tổng công suất được cung cấp cho hệ thống và $\mathbf{P} = (P_S, P_{\mathcal{R}_b})$.

Bài toán tối ưu trong (3.22) là loại bài toán không lồi vì hàm mục tiêu $\bar{\varepsilon}_{e2e}(\mathbf{P})$ là một hàm không lồi. Do đó, chúng ta cố gắng biến đổi (3.22) thành một bài toán tối ưu lồi. Hơn nữa, theo yêu cầu uRLLC $\bar{\varepsilon}_{e2e}(\mathbf{P}) \leq 10^{-5}$ [90], tương ứng với SNR ở mức cao. Từ điều này, vấn đề tối ưu (3.22) được tính xấp xỉ bên dưới như sau:

$$\min_{\mathbf{P}_S, \mathbf{P}_{\mathcal{R}_b}} \tilde{\varepsilon}_{S\mathcal{R}_b} + \tilde{\varepsilon}_{\mathcal{R}_b\mathcal{D}} \text{ s.t. } P_S + P_{\mathcal{R}_b} \leq P_0. \quad (3.23)$$

Bằng cách thay (3.18) và (3.19) vào (3.23), bài toán tối ưu (3.23) trở thành

$$\min_{\mathbf{P}_S, \mathbf{P}_{\mathcal{R}_b}} \left(\frac{\theta \sigma^2 d_1^\eta}{P_S} \right)^{NT} + \left(\frac{\theta \sigma^2 d_2^\eta}{P_{\mathcal{R}_b}} \right)^M \frac{1}{M!} \text{ s.t. } P_S + P_{\mathcal{R}_b} \leq P_0, \quad (3.24)$$

(công thức (3.24) được chứng minh trong Phụ lục B).

với d_1 và d_2 lần lượt là các khoảng cách của các kênh truyền giữa $S \rightarrow \mathcal{R}_b$ và $\mathcal{R}_b \rightarrow \mathcal{D}$, tương ứng. Ở đây, hệ thống đơn giản hóa việc xem xét mô hình suy giảm tín hiệu truyền với độ lợi trung bình của kênh như $\Omega_{S\mathcal{R}_b} = d_1^{-\eta}$ và $\Omega_{\mathcal{R}_b\mathcal{D}} = d_2^{-\eta}$ [86], trong đó η là hệ số suy hao tín hiệu qua kênh truyền.

Do hàm mục tiêu trong (3.24) là lồi nên chúng ta áp dụng phương pháp nhân tử Lagrange cho (3.24) như sau:

$$\Delta_1(\mathbf{P}, \psi_1) = \left(\frac{\theta \sigma^2 d_1^\eta}{P_S} \right)^{NT} + \left(\frac{\theta \sigma^2 d_2^\eta}{P_{\mathcal{R}_b}} \right)^M \frac{1}{M!} - \psi_1 (P_S + P_{\mathcal{R}_b} - P_0), \quad (3.25)$$

với ψ_1 là hằng số nhân tử Lagrange. Bằng cách lấy đạo hàm riêng của $\Delta_1(\mathbf{P}, \psi_1)$ theo từng biến P_S , $P_{\mathcal{R}_b}$ và ψ_1 , chúng ta nhận được hệ phương trình như sau:

$$\begin{cases} \frac{\partial \Delta_1(\mathbf{P}, \psi_1)}{\partial P_S} = 0 \\ \frac{\partial \Delta_1(\mathbf{P}, \psi_1)}{\partial P_{\mathcal{R}_b}} = 0 \\ \frac{\partial \Delta_1(\mathbf{P}, \psi_1)}{\partial \psi_1} = 0 \end{cases} \Leftrightarrow \begin{cases} \frac{-(\theta \sigma^2 d_1^\eta)^{NT} NT}{P_S^{NT+1}} - \psi_1 = 0, & (\mathcal{F}_1), \\ \frac{-(\theta \sigma^2 d_2^\eta)^M M}{P_{\mathcal{R}_b}^{M+1} M!} - \psi_1 = 0, & (\mathcal{F}_2), \\ P_S + P_{\mathcal{R}_b} - P_0 = 0, & (\mathcal{F}_3). \end{cases} \quad (3.26)$$

Bằng cách trừ $(\mathcal{F}_1) - (\mathcal{F}_2)$, $P_{\mathcal{R}_b}$ thu được sau đây:

$$P_{\mathcal{R}_b} = \left(\frac{(\sigma^2 \theta d_2^\eta)^M M}{(\sigma^2 \theta d_1^\eta)^{NT} M! NT} \right)^{1/M+1} P_S^{\frac{NT+1}{M+1}}. \quad (3.27)$$

Bằng cách thay (3.27) vào $(\mathcal{F}_3)$, chúng ta được phương trình như sau:

$$\underbrace{P_S + \mathcal{Z}_1 P_S^{\frac{NT+1}{M+1}}}_{\triangleq g(P_S)} = P_0, \quad (3.28)$$

$$\text{với } \mathcal{Z}_1 = \left(\frac{(\sigma^2 \theta d_2^\eta)^M M}{(\sigma^2 \theta d_1^\eta)^{NT} M! NT} \right)^{1/M+1} \text{ và đặt } g(P_S) = P_S + \mathcal{Z}_1 P_S^{\frac{NT+1}{M+1}}.$$

Từ (3.28), khi $P_S > 0$, $\mathcal{Z}_1 > 0$, $(M, N, T) \in \mathbb{N}^*$ và $\frac{\partial g(P_S)}{\partial P_S} > 0$ thì $g(P_S)$ là một

hàm tăng. Hơn nữa, vế phải của (3.28) là một hằng số.

Khi $g(0) = 0$, $g(P_S) > 0$ và $P_0 > 0$, phương trình (3.28) có một nghiệm duy nhất.

Để tìm nghiệm này, chúng ta dựa vào thuật toán tìm gốc Newton–Raphson được minh họa qua Bảng 3.1. Thuật toán tìm gốc Newton–Raphson như sau [91, CT. (2.31)]:

Giả sử x_0 là giá trị nghiệm khởi đầu được chọn của hàm $f(x)$, phép lặp tìm nghiệm xấp xỉ gốc duy nhất x_{n+1} của $f(x)$ được xác định như sau:

$$x_{n+1} = x_n - \frac{f(x_n)}{\partial f(x_n) / \partial x_n}. \quad (3.29)$$

(công thức (3.29) được chứng minh trong Phụ lục C)

Lưu ý rằng hàm phi tuyến tính (3.28) với các ràng buộc thỏa mãn điều kiện hội tụ cho cả phương pháp Newton-Raphson và các phương pháp khác (như Bisection, Secant, Bairstow). Tuy nhiên, Newton-Raphson có tốc độ hội tụ nhanh nhất, trong khi Bisection có tốc độ hội tụ chậm nhất với tốc độ tuyến tính [92]. Secant và Bairstow có tốc độ hội tụ số học và bậc hai, lần lượt cao hơn Bisection nhưng thấp hơn Newton-Raphson. Vì vậy, Newton-Raphson là lựa chọn tối ưu trong trường hợp này.

Dựa vào thuật toán 1 và $(\mathcal{F}_3)$, giá trị công suất tối ưu của nút $\mathcal{R}_b$ tính được như sau:

$$P_{\mathcal{R}_b}^* = P_0 - P_S^*. \quad (3.30)$$

Trong trường hợp đặc biệt với $NT = M$, phương trình (3.28) được biến đổi thành một phương trình bậc nhất như sau:

$$(1 + \mathcal{Z}_1) P_S = P_0. \quad (3.31)$$

Thay $\mathcal{Z}_1$ vào (3.31) và kết hợp $(\mathcal{F}_3)$, công suất phát tối ưu ở nút $\mathcal{S}$ và nút $\mathcal{R}_b$ như sau:

$$\begin{cases} P_S^* = \phi P_0, \\ P_{\mathcal{R}_b}^* = (1 - \phi) P_0, \end{cases} \quad (3.32)$$

$$\text{với } \phi = \left(1 + \left(\frac{1}{M!} \right)^{1/M+1} \left(\frac{d_2}{d_1} \right)^{\frac{\eta M}{M+1}} \right)^{-1}.$$

3.4.2 Tối ưu vị trí nút chuyển tiếp

Ở mục này, chúng ta sẽ giải quyết vấn đề tìm vị trí tối ưu cho nút chuyển tiếp trong hệ thống, tức là tìm khoảng cách tối ưu đặt $\mathcal{R}_b$ nhằm giảm thiểu BLER toàn trình của hệ thống, đồng thời có thể đáp ứng được yêu cầu của uRLLCs. Quá trình này sẽ được thực hiện trong bối cảnh tổng công suất phát đã được xác định trước và khoảng cách truyền từ đầu $\mathcal{S}$ đến cuối $\mathcal{D}$ cũng đã được cho trước. Hệ thống mạng được đề xuất chủ yếu hướng tới ứng dụng trong các mạng di động 5G và IoT, nơi các thiết bị đầu cuối có thể linh hoạt điều chỉnh để tương thích với mạng mục tiêu. Đồng thời, hệ thống mạng được nghiên cứu trong khoảng thời gian rất ngắn, do đó các nút mạng tham gia chủ yếu là tĩnh. Cụ thể, vấn đề tối ưu vị trí được mô tả như sau:

$$\min_{\mathbf{d}} \bar{\varepsilon}_{e2e}(\mathbf{d}) = \min_{d_1, d_2} \tilde{\varepsilon}_{\mathcal{S}\mathcal{R}_b} + \tilde{\varepsilon}_{\mathcal{R}_b\mathcal{D}} \text{ s.t. } d_1 + d_2 = D. \quad (3.33)$$

với $\mathbf{d} = (d_1, d_2)$ và D là khoảng cách tổng cộng từ nút nguồn đến nút đích.

Bảng 3.1: Thuật toán 1 tìm nghiệm gốc của (3.28).

Thuật toán 1	Phép lặp tìm nghiệm (3.28)
Đầu vào	$\theta, N, M, d_1^\eta, d_2^\eta, \sigma^2, \eta, T, P_0$
Đầu ra	P_S^*
Các bước:	Khởi tạo số lần lặp tối đa $N_{\max} = 1000$ và 1) $P_{S,0} = 1$ 2) Tính: $\mathcal{Z}_1 = \left(\frac{(\sigma^2 \theta d_2^\eta)^M M}{(\sigma^2 \theta d_1^\eta)^{NT} M! NT} \right)^{1/M+1}$ 3) Tạo vòng lặp: for $0 \leq n \leq N_{\max}$ do 4) $f(P_{S,n}) = P_{S,n} + \mathcal{Z}_1 P_{S,n}^{\frac{NT+1}{M+1}} - P_0$ 5) $\frac{\partial f(P_{S,n})}{\partial P_{S,n}} = 1 + \frac{N+1}{M+1} \mathcal{Z}_1 P_{S,n}^{\frac{NT-M}{M+1}}$

$$\begin{array}{l}
6) \quad P_{S,n+1} = P_{S,n} - \frac{f(P_{S,n})}{\partial f(P_{S,n}) / \partial P_{S,n}}. \\
7) \quad \text{Kết thúc vòng lặp for} \\
8) \quad \text{Tìm được: } P_S^* = P_{S,N_{\max}+1}
\end{array}$$

Bằng cách thay thế (3.18) và (3.19) vào (3.33), vấn đề tối ưu có thể được viết lại như sau:

$$\min_{\mathbf{d}} \left(\frac{\theta \sigma^2 d_1^\eta}{P_S} \right)^{NT} + \left(\frac{\theta \sigma^2 d_2^\eta}{P_{\mathcal{R}_b}} \right)^M \frac{1}{M!} \text{ s.t. } d_1 + d_2 = D. \quad (3.34)$$

Bài toán trong (3.34) là một bài toán tối ưu lồi. Do đó, tương tự như phần tối ưu công suất, phương pháp nhân tử Lagrange được lựa chọn để tìm vị trí tối ưu tổng quát $\mathbf{d}^*$. Hàm Lagrange cho (3.34) được viết lại như sau:

$$\Delta_2(\mathbf{d}, \psi_2) = \left(\frac{\theta \sigma^2 d_1^\eta}{P_S} \right)^{NT} + \left(\frac{\theta \sigma^2 d_2^\eta}{P_{\mathcal{R}_b}} \right)^M \frac{1}{M!} - \psi_2 (d_1 + d_2 - D), \quad (3.35)$$

với ψ_2 là hằng số nhân tử Lagrange.

Bằng cách lấy đạo hàm riêng của $\Delta_2(\mathbf{d}, \psi_2)$ lần lượt theo các biến d_1 , d_2 và ψ_2 , bằng không. Hệ phương trình thu được bên dưới như sau:

$$\begin{cases} \frac{\partial \Delta_2(\mathbf{d}, \psi_2)}{\partial d_1} = 0 \\ \frac{\partial \Delta_2(\mathbf{d}, \psi_2)}{\partial d_2} = 0 \\ \frac{\partial \Delta_2(\mathbf{d}, \psi_2)}{\partial \psi_2} = 0 \end{cases} \Leftrightarrow \begin{cases} \frac{(\theta \sigma^2)^{NT}}{P_S^{NT+1}} NT \eta d_1^{N\eta-1} - \psi_2 = 0, & (\mathcal{G}_1), \\ -\frac{(\theta \sigma^2)^M M \eta d_2^{N\eta-1}}{P_{\mathcal{R}_b}^{M+1} M!} - \psi_2 = 0, & (\mathcal{G}_2), \\ d_1 + d_2 - D = 0, & (\mathcal{G}_3). \end{cases} \quad (3.36)$$

Lấy $(\mathcal{G}_1) - (\mathcal{G}_2)$, chúng ta có kết quả như sau:

$$d_2^{\eta M-1} = \left(\frac{(\theta \sigma^2)^{NT} NTP_{\mathcal{R}_b}^M M!}{P_S^{NT} (\theta \sigma^2)^M M} \right)^{\frac{1}{M\eta-1}} d_1^{\frac{NT\eta-1}{M\eta-1}}. \quad (3.37)$$

Thay (3.37) vào $(\mathcal{G}_3)$, phương trình của d_1 được dẫn ra như bên dưới như sau:

$$d_1 + \mathcal{Z}_2 d_1^{\frac{NT\eta-1}{M\eta-1}} = D, \quad (3.38)$$

$$\text{với } \mathcal{Z}_2 = \left(\frac{(\sigma^2 \theta)^N NTP_{\mathcal{R}_b}^M M!}{(\sigma^2 \theta)^M MP_S^{NT}} \right)^{1/M\eta-1}.$$

Tương tự như (3.28), phương trình trong (3.38) cũng có một nghiệm duy nhất d_1^* . Chúng ta áp dụng thuật toán tìm gốc Newton–Raphson để tìm nghiệm cho (3.38), và quy trình tìm nghiệm này được mô tả như trong Bảng 3.2.

Bằng cách thay giá trị tối ưu d_1^* tìm được vào $(\mathcal{G}_3)$ thu được d_2^* như sau:

$$d_2^* = D - d_1^*. \quad (3.39)$$

Đối với trường hợp đặc biệt với $NT = M$, thay $\mathcal{Z}_2$ vào (3.38) và kết hợp $(\mathcal{G}_3)$, các khoảng cách tối ưu để đặt $\mathcal{R}_b$ được xác định là:

$$\begin{cases} d_1^* = D \left(1 + (M!)^{1/\eta M-1} \left(\frac{P_{\mathcal{R}_b}}{P_S} \right)^{\frac{M}{\eta M-1}} \right)^{-1}, \\ d_2^* = D - d_1^*. \end{cases} \quad (3.40)$$

Bảng 3.2: Thuật toán 2 tìm nghiệm gốc của (3.38).

Thuật toán 2	Phép lặp tìm nghiệm (3.38)
Đầu vào	$\theta, N, M, P_S, P_{\mathcal{R}}, \sigma^2, \eta, T, D$
Đầu ra	d_1^*
Các bước:	

- 1) Khởi tạo số lần lặp tối đa $N_{\max} = 1000$ và $d_{1,0} = 1$
- 2) Tính $\mathcal{Z}_2 = \left(\frac{(\sigma^2 \theta)^{NT} NTP_{\mathcal{R}}^M M!}{(\sigma^2 \theta)^M MP_S^{NT}} \right)^{1/M\eta-1}$
- 3) Tạo vòng lặp: for $0 \leq n \leq N_{\max}$ do
- 4) $f(d_{1,n}) = d_{1,n} + \mathcal{Z}_2 d_{1,n}^{\frac{NT\eta-1}{M\eta-1}} - D$
- 5) $\frac{\partial f(d_{1,n})}{\partial d_{1,n}} = 1 + \frac{NT\eta-1}{M\eta-1} \mathcal{Z}_2 d_{1,n}^{\frac{\eta(NT-M)}{M\eta-1}}$
- 6) $d_{1,n+1} = d_{1,n} - \frac{f(d_{1,n})}{\partial f(d_{1,n}) / \partial d_{1,n}}$
- 7) Kết thúc vòng lặp for
- 8) Tìm được $d_1^* = d_{1,N_{\max}+1}$

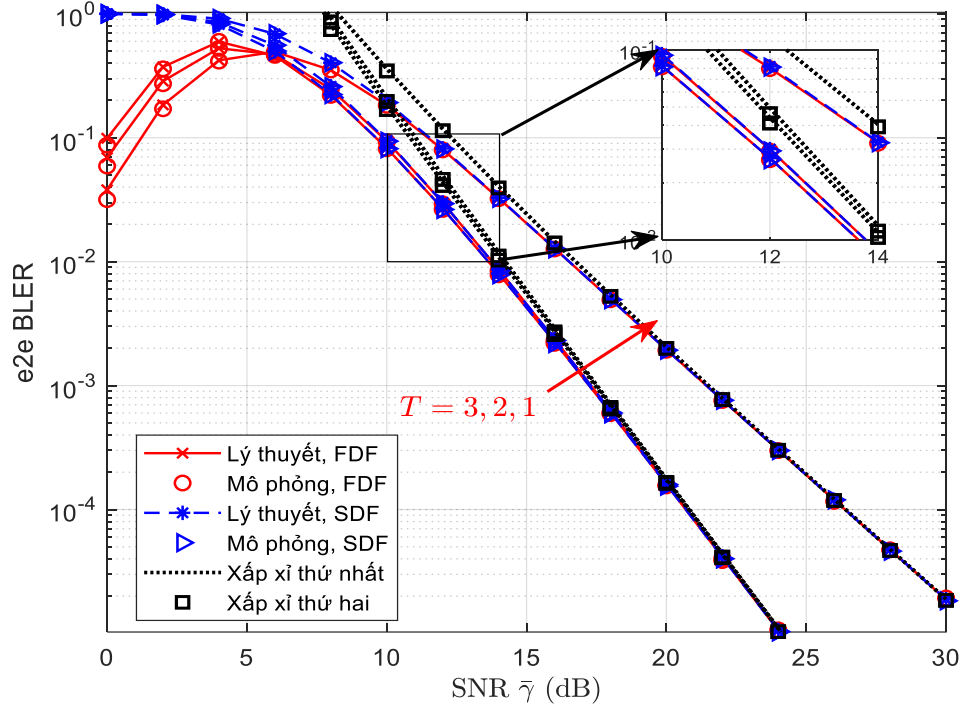
3.5 Mô phỏng và thảo luận

Ở đây, chúng ta sử dụng phương pháp mô phỏng Monte Carlo để xác minh các nội dung mới được đưa ra ở phần trình bày lý thuyết bên dưới như sau:

A. Đánh giá hiệu năng hệ thống

Trước tiên, chúng ta thiết lập các tham số mô phỏng được xác định như sau: số bit truyền tin là $\beta = 256$ bits [19] và chiều dài của khối tin FBL là $k = 256$ CUs; tổng khoảng cách truyền tin chuẩn hóa từ đầu $\mathcal{S}$ đến cuối $\mathcal{D}$ là $D = 10$; khoảng cách kênh truyền lần lượt từ $\mathcal{S} \rightarrow \mathcal{R}_b$ và $\mathcal{R}_b \rightarrow \mathcal{D}$ là $d_1 = 0.7D$ và $d_2 = 0.3D$. Hơn nữa, hệ số suy giảm tín hiệu truyền tin tương ứng là $\eta = 3$ [93]. Chúng ta chú ý rằng $\bar{\gamma} = \frac{P_0}{\sigma^2}$ là SNR trung bình của nguồn phát với tổng công suất được phân chia cho hệ thống P_0 . Bởi vì σ^2 là một hằng số, việc tăng SNR cũng đồng nghĩa với việc tăng công suất của nguồn phát. Trong hệ thống này, chúng ta giả định rằng sự phân chia công suất không đồng đều cho các nút $\mathcal{S}$ và $\mathcal{R}_b$ như $P_S = 0.7P_0$ và

$P_{\mathcal{R}_b} = 0.3P_0$. Trước hết, chúng ta tiến hành so sánh BLER toàn trình trong hai tình huống: sử dụng kỹ thuật FDF, và SDF. BLER toàn trình được nghiên cứu như một hàm của SNR trung bình $\bar{\gamma}$, dựa trên giả định số nút của $\mathcal{R}_n$ là $N = 2$ và số anten ở nút $\mathcal{D}$ là $M = 3$, như được mô tả trong Hình 3.2..

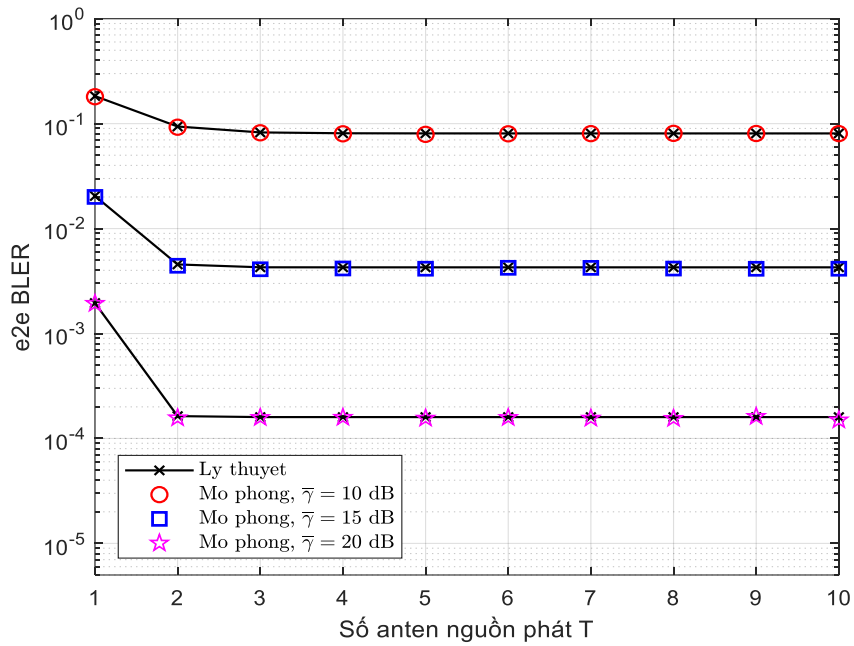


Hình 3.2: BLER toàn trình của SDF và FDF thay đổi theo SNR với $N = 2$ và $M = 3$.

Quan sát Hình 3.2, BLER toàn trình của SDF và FDF giống nhau ở vùng SNR trung bình và cao. Tuy nhiên, ở vùng SNR thấp ($\bar{\gamma} < 6\text{dB}$) hiệu năng của FDF tốt hơn so với SDF do BLER của FDF nhỏ hơn.

Do đó, kỹ thuật FDF được chọn để khảo sát trong các hình vẽ ở phần sau. Hơn nữa, sự tác động của số lượng anten ở nút nguồn T vào BLER toàn trình của hệ thống cũng được mô tả với các giá trị $T = 1, 2, 3$. Hiệu năng được cải tiến đáng kể khi T tăng từ 2 đến 3 ở vùng SNR trung bình và thấp ($\bar{\gamma} < 15\text{dB}$). Tuy nhiên, khi $\bar{\gamma}$ đủ lớn hiệu năng hầu như không thay đổi (BLER là hoàn toàn giống nhau)

ở các giá trị $T = 2$ và $T = 3$. Do vậy, việc chọn giá trị T cần phải xem xét để thỏa mãn các yêu cầu cải thiện hiệu năng và hạn chế được chi phí lắp đặt hệ thống. Tiếp theo, Hình 3.3 xét ảnh hưởng của T vào BLER toàn trình, trong đó $N = 2$, $M = 3$ và $\bar{\gamma} = 10, 15, 20$ dB. Như quan sát trong Hình 3.3 BLER bảo hòa khi $T = 2$. Do đó, $T = 2$ là một lựa chọn hiệu quả để cân nhắc giữa hiệu năng và chi phí thực hiện. Một cách để giảm BLER của hệ thống như mong muốn là tăng công suất nguồn phát.

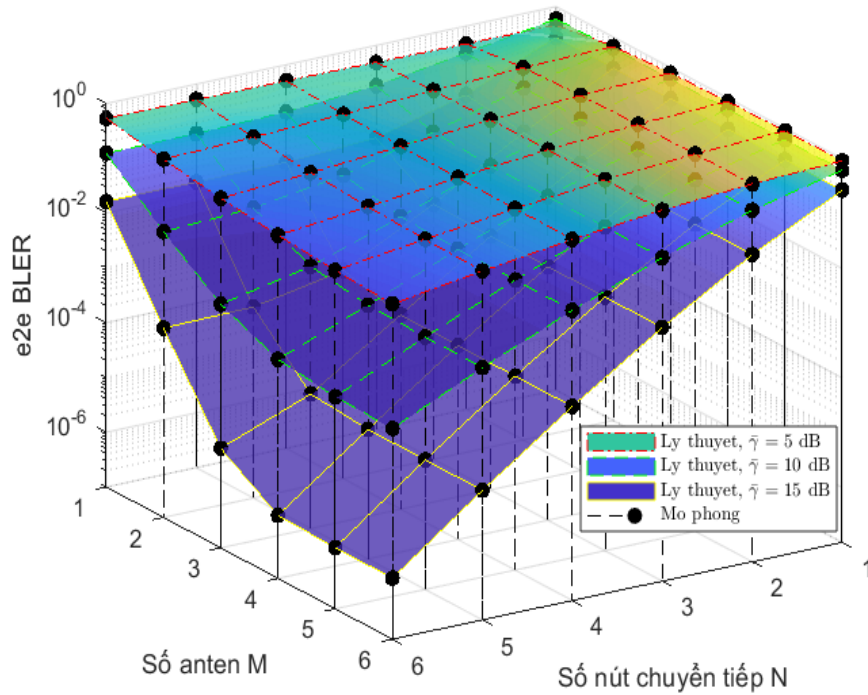


Hình 3.3: Sự ảnh hưởng e2e BLER vào số an-ten nguồn phát T với $N = 2$ và $M = 3$.

Tuy nhiên, công suất phát thường bị hạn chế trong thực tế, và thậm chí khi có khả năng tăng công suất phát, điều này dẫn đến nhiễu trong các kênh truyền khác. Trong khi đó, tăng số nút N và số an-ten M tại nút đích có thể xem xét như là một lựa chọn thay thế nhằm nâng cao hiệu năng cho hệ thống, như minh họa trong Hình 3.4. Hình này phân tích e2e BLER phụ thuộc vào số nút N và số an-ten M ở nút đích trong không gian ba chiều. Mỗi mặt phẳng trình bày kết quả phân tích BLER với SNR trung bình $\bar{\gamma}$ tương ứng với ba mức $\bar{\gamma} = 5, 10$ dB và 15dB với

$T = 2$. Ta thấy được rằng đường mô phỏng (dấu chấm) trùng khớp với mặt phẳng kết quả lý thuyết.

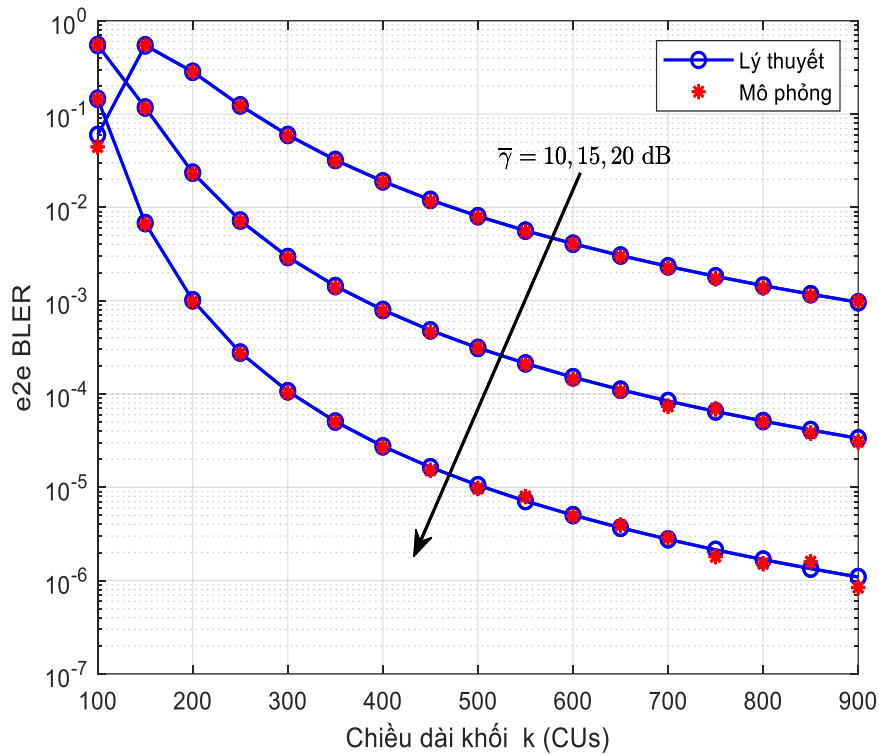
Quan sát Hình 3.4, e2e BLER giảm đáng kể khi $\bar{\gamma}$, N và M tăng. Chúng ta để ý rằng tăng độ phân tập của hệ thống phụ thuộc vào N , M và T và phù hợp với kết quả trong (3.21). Do đó, điều này làm nổi bật tầm quan trọng của việc thiết kế số nút N và số an-ten một cách chính xác sao cho hiệu năng được cải thiện như dự kiến và đồng thời chi phí lắp đặt được giảm xuống. Cụ thể, e2e BLER $\approx 10^{-5}$ ở $\bar{\gamma} = 15$ dB và $T = 2$, số nút chuyển tiếp và số an-ten có thể chọn $M = 3$ và $N = 4$. Do vậy, chúng ta cũng chọn giá trị $M = 3$ và $N = 4$ cho các mô phỏng từ Hình 3.4 trở đi.



Hình 3.4: e2e BLER theo số nút N và số an-ten M với ba trường hợp $\bar{\gamma} = 10, 15$ dB và 20 dB.

Tiếp đến, Hình 3.5 biểu diễn ảnh hưởng của k đối với BLER toàn trình trong ba tình huống SNR như $\bar{\gamma} = 10, 15$ dB và 20 dB với $T = 2$, $N = 4$ và $M = 3$

tương ứng. Trong Hình 3.5, hiệu năng hệ thống được cải thiện bằng cách tăng SNR trung bình và chiều dài khối k . Tuy nhiên, chiều dài khối lớn cũng gây ra độ trễ toàn trình cao hơn, như được mô tả trong Hình 3.7. Do đó, khi cải thiện hiệu năng dựa vào việc tăng SNR hay tăng chiều dài khối cần phải thận trọng. Ví dụ, đối với $\bar{\gamma} = 20$ dB, để đảm bảo yêu cầu về độ tin cậy theo yêu cầu BLER $\approx 10^{-5}$ có thể chọn chiều dài khối 450 CUs.

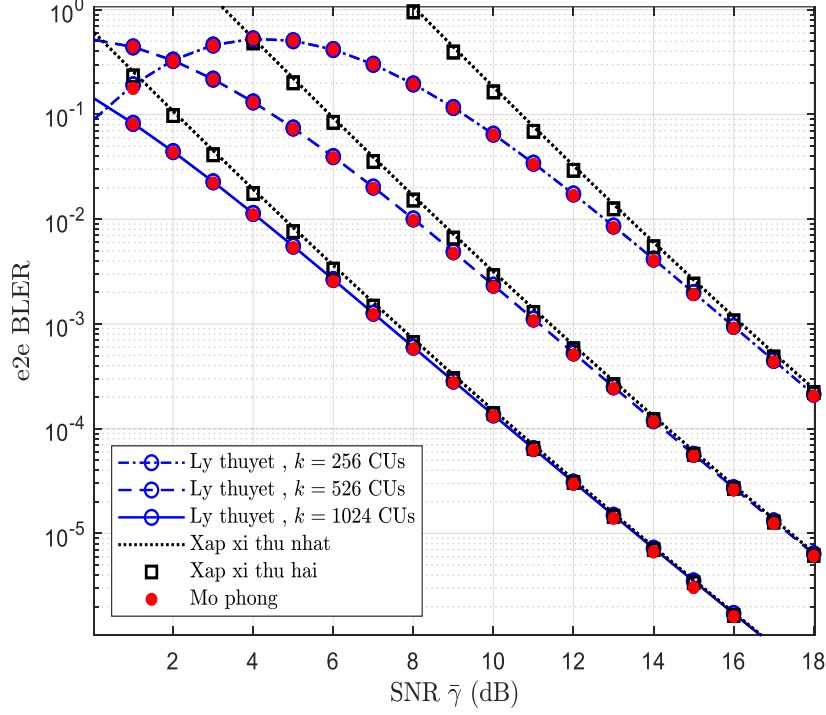


Hình 3.5: $e2e$ BLER theo độ dài khối tin k với ba trường hợp $\bar{\gamma} = 10, 15$ dB và 20 dB.

Kế tiếp, Hình 3.6 mô tả sự ảnh hưởng của giá trị $e2e$ BLER gần đúng và tiệm cận vào $\bar{\gamma}$ trong trường hợp $T = 2$, $N = 4$, $M = 3$ và $k = 256, 512$ CUs và 1024 CUs, tương ứng.

, BLER giảm khi chúng ta tăng $\bar{\gamma}$ và k . Các đường cong của kết quả lý thuyết và đường mô phỏng hội tụ với các đường tiệm cận ở vùng $\bar{\gamma}$ cao. Hơn nữa, đường

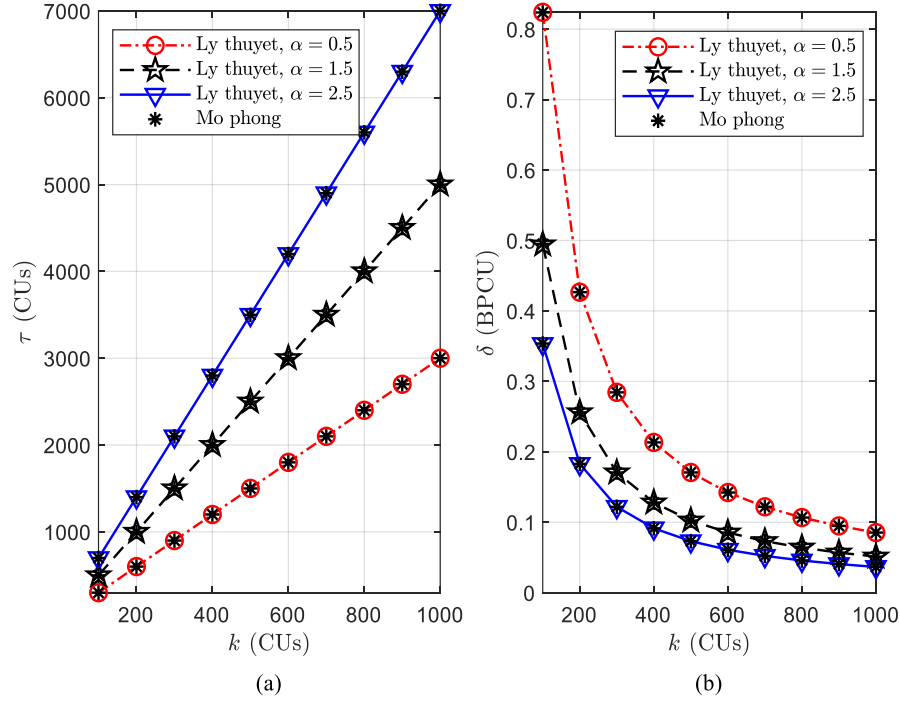
tiệm cận thứ hai hoàn toàn khớp với đường tiệm cận thứ nhất và được coi là một cách đánh giá hiệu quả của đường tiệm thứ nhất như trong (3.20).



Hình 3.6: Sự phụ thuộc e2e BLER gần đúng và tiệm cận vào $\bar{\gamma}$ trong các trường hợp $k = 256, 512$ CUs và 1024 CUs.

Cuối cùng, Hình 3.7 minh họa sự phụ thuộc của độ trễ toàn trình và thông lượng vào chiều dài khối k với $T=2$, $N=4$, $M=3$, $\bar{\gamma}=20$ dB và $\alpha=0.5, 1.5$ và 2.5 , tương ứng. Đây cũng là các thông số đánh giá quan trọng theo yêu cầu uRLLCs. Khi tăng chiều dài khối k thì độ tin cậy tăng, tuy nhiên, độ trễ toàn trình tăng và thông lượng của hệ thống giảm. Bên cạnh đó, Hình 3.7 cũng cho thấy tác động của hệ số trễ giải mã. Cụ thể, giá trị α càng tăng thì độ trễ toàn trình càng tăng và thông lượng càng giảm. Để ý rằng, nếu chúng ta chọn độ dài khối là $k=450$ CUs, hệ thống sẽ bị trễ khoảng 2250 CUs tương ứng với 6.75 ms (bởi vì thời gian 1 CUs tương ứng với 3 μ s [94]). Độ trễ này hoàn toàn đáp ứng yêu cầu

về giới hạn độ trễ thấp của uRLLC để hỗ trợ các ứng dụng hoạt động trong môi trường tại nhà máy [94].

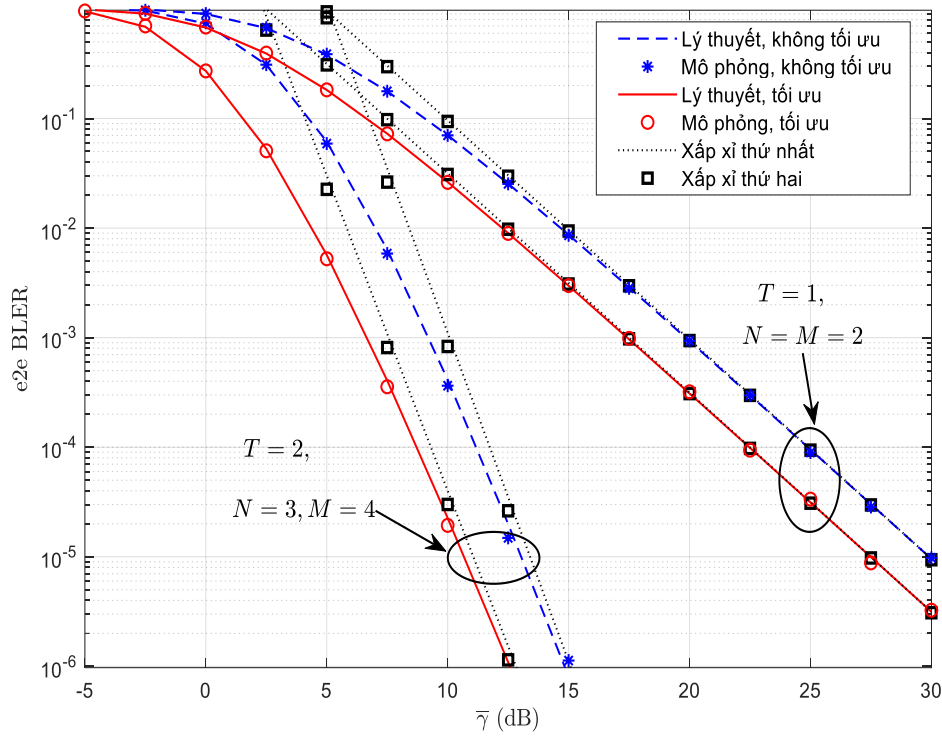


Hình 3.7: Độ trễ và thông lượng toàn trình theo chiều dài khối k ở $\bar{\gamma} = 20\text{dB}$ với các trường hợp $\alpha = 0.5, 1.5$ và 2.5 .

B. Đánh giá hiệu năng trường hợp tối ưu công suất và tối ưu vị trí

Trong phần này, chúng ta so sánh hiệu năng trong trường hợp không tối ưu với trường hợp tối ưu, đối với cả công suất và vị trí nút chuyển tiếp. Để bảo đảm sự công bằng, chúng ta áp dụng các tham số mô phỏng tương tự như Phần 3.5A cho trường hợp không tối ưu, chi tiết như sau: $\beta = 256$ bits và độ dài của khối tin $k = 256$ CUs; hệ số suy giảm đường truyền $\eta = 3$; tổng khoảng cách truyền được chuẩn hóa với $D = 10$. Trường hợp tổng quát, các số liệu cụ thể của T , N và M được thiết lập như $T = 2$, $N = 3$ và $M = 4$. Trường hợp đặc biệt, các giá trị $T = 1$, $N = M = 2$ được chọn. Hình 3.8 so sánh hiệu năng hệ thống giữa cấu hình tối ưu và không tối ưu công suất trong cả hai tình huống tổng quát và đặc biệt của N và M . Trong đó, khoảng cách được thiết lập cố định như giả thuyết ở Phần A với

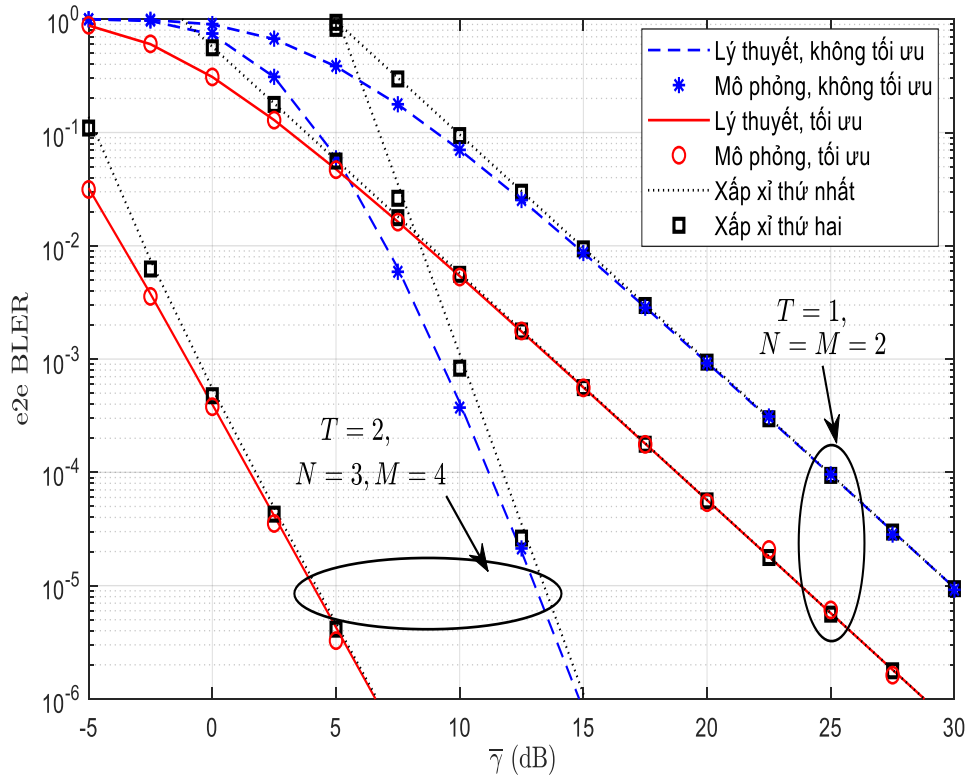
$d_1 = 0.7D$ và $d_2 = 0.3D$. Để làm nổi bật hiệu năng vượt trội của cấu hình tối ưu, chúng ta xem xét cấu hình không tối ưu với công suất bằng nhau, tức là $P_S = P_{R_b} = 0.5P_0$.



Hình 3.8: $e2e$ BLER theo SNR trung bình giữa trường hợp tối ưu và không tối ưu công suất.

Hình 3.8 cho thấy rằng hiệu năng trường hợp tối ưu công suất cao hơn đáng kể so với trường hợp không tối ưu ở mọi giá trị SNR cho tình huống tổng quát và đặc biệt. Như quan sát Hình 3.8, với $BLER \approx 10^{-5}$ độ lợi mảng (đoạn chênh lệch SNR giữa đường tối ưu và không tối ưu) đạt được là 2.5dB cho tất cả trường hợp $T=1$, $N=M=2$ và $T=2$, $N=3$ và $M=4$. Tiếp theo, Hình 3.9 minh họa hiệu năng hệ thống trong tình huống tối ưu vị trí và không tối ưu vị trí với $d_1 = d_2 = 0.5D$. Trong hình này, chúng ta xem xét công suất phát là cố định với $P_S = 0.7P_0$ và $P_{R_b} = 0.3P_0$, như đã giả định trong Phần 3.5A. Như quan sát Hình 3.9, hiệu năng

trong trường hợp tối ưu vị trí vượt trội hơn hẳn trường hợp không tối ưu tương tự như ở Hình 3.8. Tuy nhiên, độ lợi mảng đạt được trong Hình 3.9 cao hơn nhiều so với Hình 3.8. Cụ thể, đối với $BLER \approx 10^{-5}$, độ lợi mảng thu được là 6.25dB và 9.5dB cho các trường hợp $T=1, N=M=2$ và $T=2, N=3$ và $M=4$, tương ứng.



Hình 3.9: $e2e$ BLER theo SNR trung bình trong tình huống giữa tối ưu và không tối ưu vị trí.

3.6 Kết luận Chương 3

Chương 3 đã đưa ra một mô hình mạng DF hai chặng mới trong SPC. Mô hình mạng đề xuất này mang lại lợi ích như sau: i) Phù hợp cho hệ thống cần phải hạn chế công suất nguồn phát và giảm nhiễu đồng kênh; ii) Nâng cao khả năng phân tập không gian của hệ thống SPC bằng cách kết hợp các kỹ thuật TAS, PRS và MRC tại các thiết bị đầu cuối tương ứng.

Chương 3 cũng đề xuất các phương pháp cho phân tích hiệu năng của hệ thống, cụ thể là phân tích BLER dạng chính xác và dạng xấp xỉ cho các phương pháp SDF và FDF qua kênh truyền fading Rayleigh. Tiếp theo, độ trễ và thông lượng toàn trình của hệ thống cũng được phân tích. Cuối cùng, kết quả mô phỏng đã làm rõ độ chính xác của kết luận lý thuyết, và cho thấy hệ thống đã đạt đến mức phân tập hoàn chỉnh. Điểm đặc biệt mới trong Chương 3 này là cách tiếp cận tối ưu hệ thống, cụ thể thông qua hai chiến lược quan trọng: tối ưu công suất nguồn phát và tối ưu vị trí của nút chuyển tiếp. Đồng thời các chiến lược này tuân theo các ràng buộc như tổng công suất phân bổ cho hệ thống và khoảng cách không đổi giữa nút đầu cuối. Đường mô phỏng (đánh dấu tròn) đã minh chứng rằng chiến lược tối ưu về công suất và vị trí đều vượt trội hơn đáng kể so với tình huống không tối ưu.

Mô hình hệ thống trong chương 3 được thiết kế đặc biệt cho mạng 5G, với các đặc điểm như công suất nguồn phát hạn chế, vùng phủ sóng hẹp, khoảng cách truyền ngắn và nhiều đồng kênh lớn. Tuy nhiên, mô hình này cũng có một số hạn chế trong việc thực thi hệ thống, như chi phí cao do thiết kế đa an-ten ở nguồn phát, nhiều nút chuyển tiếp và nhiều an-ten thu.

Cuối cùng, tất cả nội dung của Chương 3 được minh chứng trong kết quả công bố số 2 của NCS, đó là: N. T. Y. Linh, T. Ngo Hoang, P. N. Son, and V. N. Q. Bao, "Dual-hop Relaying Networks for Short-Packet URLLCs: Performance analysis and Optimization," *Journal of Communications and Networks*, vol. 24, no. 4, pp. 408-418, Aug. 2022, doi: 10.23919/JCN.2022.000020. **(SCIE-Q1)**

Chương 4. PHÂN TÍCH HIỆU NĂNG HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG CHUYỂN TIẾP SỬ DỤNG MẶT PHẢN XẠ THÔNG MINH

4.1 Giới thiệu

Mặc dù, hệ thống truyền SPC trong mạng DF hai chặng ở Chương 3 đã góp phần hiện thực cho giao tiếp uRLLCs bằng cách nâng cao độ phân tập cho hệ thống và tối ưu. Tuy nhiên, sử dụng nhiều an-ten tại nguồn phát, nguồn nhận và nút trung gian đã dẫn đến năng lượng tiêu thụ lớn. Hơn nữa, hoạt động của hệ thống diễn ra trong hai khoảng thời gian liên tiếp nên việc đảm bảo yêu cầu nghiêm ngặt về độ trễ trở nên khó khăn. Ngoài ra, cho đến thời điểm nghiên cứu của Luận án, mặt phản xạ IRS đã được xác định là một trong những giải pháp tối ưu nhằm giảm thiểu tiêu thụ điện năng trong quá trình truyền dẫn, đồng thời nâng cao hiệu quả sử dụng phổ tần trong thiết kế hệ thống mạng. Cấu tạo, nguyên tắc hoạt động và các ưu điểm của IRS đã được trình bày chi tiết trong Chương 2 của Luận án.

Hơn nữa, một số công trình nghiên cứu đã chỉ ra rằng truyền thông tin qua IRS không chỉ cải thiện hiệu năng hệ thống mà còn giúp giảm công suất phát của nguồn phát so với truyền thông tin qua nút chuyển tiếp [95]. Cụ thể, nghiên cứu trong [96] đã chứng minh rằng hệ thống có sự hỗ trợ của IRS cải thiện hiệu năng so với hệ thống chuyển tiếp AF. Càng tăng yếu tố phản xạ của IRS, xác suất dừng (OP) của hệ thống sẽ giảm đáng kể so với OP của hệ thống chuyển tiếp AF. Nghiên cứu trong [31] cho thấy rằng khi IRS được thiết kế với nhiều yếu tố phản xạ, hiệu quả năng lượng và công suất truyền sẽ được cải thiện rõ rệt so với các hệ thống truyền tin trực tiếp và qua nút chuyển tiếp DF. Nghiên cứu trong [95] cũng chỉ ra rằng việc sử dụng IRS giúp mở rộng vùng phủ sóng và giảm độ trễ so với các hệ thống chuyển tiếp qua nút chuyển tiếp DF và truyền trực tiếp.

Tuy nhiên, các nghiên cứu này chủ yếu tập trung vào việc so sánh sự cải thiện của truyền thông tin gói tin dài qua IRS với truyền tin gói tin dài qua nút chuyển tiếp cơ bản. Theo hiểu biết của chúng tôi, chưa có công trình nghiên cứu nào quan

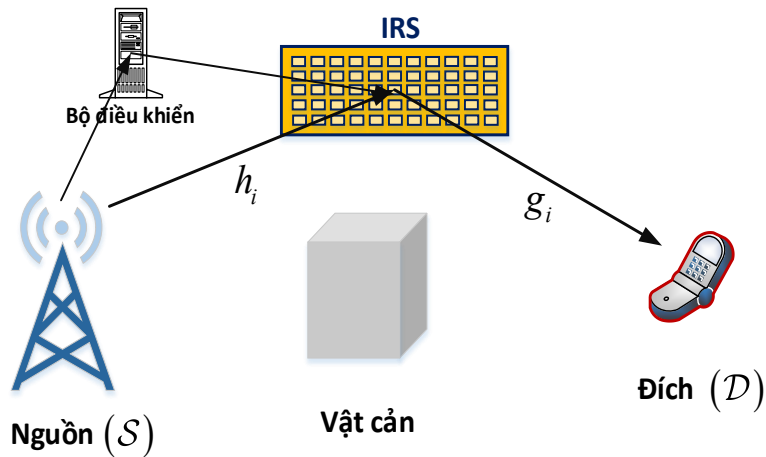
tâm đến SPC qua IRS, cũng như việc so sánh hiệu năng giữa hai mô hình SPC này. Do đó, để đáp ứng yêu cầu của uRLLCs trong các mạng thế hệ mới, Luận án đề xuất mô hình SPC qua IRS, đồng thời phân tích và so sánh hệ thống giao tiếp này với hệ thống giao tiếp SPC qua nút chuyển tiếp DF và AF.

Cấu trúc của Chương 4 được sắp xếp theo trình tự như sau: Mục 4.2 trình bày mô hình mới của hệ thống SPC sử dụng IRS và mô hình SPC trong mạng chuyển tiếp DF và AF cơ bản. Sau đó, Mục 4.3 giới thiệu các phương pháp để tính các tham số hiệu năng cho cả hai hệ thống qua kênh truyền fading Rayleigh như BLER dạng chính xác và xấp xỉ. Mục 4.4 đi sâu vào phân tích các kết luận mô phỏng, làm rõ ưu điểm của hệ thống SPC khi được hỗ trợ bởi IRS. Cuối cùng, Mục 4.5 trình bày ngắn gọn về kết luận đạt được và những nội dung mới của Chương 4.

4.2 Mô hình hệ thống

4.2.1 Hệ thống chuyển tiếp sử dụng mặt phản xạ thông minh

Hình 4.1 là mô hình hệ thống SPC qua mặt IRS, bao gồm một nút nguồn (S), một nút đích (D) (đóng vai trò nguồn nhận) và một tấm IRS. Tấm IRS này bao gồm L yếu tố phản xạ thụ động. Trong hệ thống này, S và D được giả định dùng một an-ten để truyền và nhận tín hiệu. Hệ thống không xét kênh truyền trực tiếp giữa S và D do vật cản xung quanh như tòa nhà, cây xanh. Kênh truyền giữa S và D được mô phỏng như là kênh fading Rayleigh bán tĩnh phẳng.



Hình 4.1: Mô hình SPC được trợ giúp của IRS.

Hơn nữa, hoạt động của IRS được mô tả như một hệ thống phản xạ thông minh, có khả năng phối hợp và trao đổi CSI từ cả hai nút $\mathcal{S}$ và $\mathcal{D}$ qua bộ điều khiển không dây [97]. IRS có khả năng điều chỉnh pha của các yếu tố phản xạ trong IRS sao cho SNR tại nút $\mathcal{D}$ đạt giá trị lớn nhất [27]. Do tín hiệu bị suy giảm đáng kể khi truyền trong môi trường, nên tín hiệu phản xạ lần đầu trên IRS được xem xét, trong khi các tín hiệu phản xạ từ lần thứ hai trở đi được bỏ qua [98].

Hơn nữa, hệ thống giao tiếp được xét trong trường gần, nghĩa là khoảng cách giữa nguồn hoặc đích đến chính giữa IRS nhỏ hơn $\frac{2D^2}{\lambda}$, với D là kích thước lớn nhất của an-ten và λ là bước sóng của nguồn phát sóng [71].

a. Mô hình kênh truyền:

Tín hiệu nhận tại nút $\mathcal{D}$ có dạng như sau:

$$y_{\mathcal{D}} = \sqrt{P} \left(\sum_{i=1}^L h_i r_i g_i \right) x_{\mathcal{D}} + n_{\mathcal{D}}, \quad (4.1)$$

với P là công suất của nguồn $\mathcal{S}$; $x_{\mathcal{D}}$ là tín hiệu được truyền từ nguồn $\mathcal{S}$ đến $\mathcal{D}$, $n_{\mathcal{D}}$ là nhiễu Gauss tại nút $\mathcal{D}$ với $n_{\mathcal{D}} \sim CN(0, N_0)$; r_i là hệ số phản xạ của thành phần phản xạ thứ i trong IRS ($i = 1, 2, \dots, L$) với $r_i = |r_i| e^{j\phi_i}$ [96], trong đó $|r_i|$ đại diện cho biên độ của hệ số phản xạ thứ i trong IRS (chúng ta giả định rằng IRS phản xạ hoàn toàn nên $|r_i| = 1$) và ϕ_i là góc dịch chuyển pha tương ứng.

Các hệ số kênh, bao gồm h_i và g_i , lần lượt biểu thị cho các hệ số kênh truyền từ nút $\mathcal{S}$ đến thành phần phản xạ thứ i trong IRS và từ thành phần thứ i này đến nút $\mathcal{D}$, tương ứng. Các hệ số kênh này được mô tả như sau:

$$h_i = \alpha_i e^{-j\theta_i}, \quad (4.2)$$

$$g_i = \beta_i e^{-j\varphi_i}, \quad (4.3)$$

với $\alpha_i, \beta_i, \theta_i$ và φ_i lần lượt chỉ các biên độ và góc dịch chuyển pha của các hệ số kênh truyền h_i và g_i , tương ứng. Các biên độ α_i và β_i là các biến ngẫu nhiên độc lập (RVs), có phân phối Rayleigh với trung bình $\frac{1}{2} \sqrt{\frac{\pi}{d_1^\eta}}$ và độ lệch chuẩn $\left(1 - \frac{\pi}{2}\right) \frac{1}{d_2^\eta}$; η là hệ số suy hao đường truyền; d_1 và d_2 là các khoảng cách liên kết từ $\mathcal{S} \rightarrow \text{IRS}$ và từ $\text{IRS} \rightarrow \mathcal{D}$, tương ứng.

Hơn nữa, chúng ta giả sử rằng tất cả hệ số kênh truyền đều chịu ảnh hưởng của cả hai thành phần fading lớn và nhỏ, được mô tả tổng quát như sau $h_{xy} = \sqrt{\lambda_{xy}} \tilde{h}_{xy}$, trong đó $\tilde{h}_{xy} \sim CN(0,1)$ là phần fading nhỏ, và λ_{xy} là phần fading lớn, phụ thuộc vào khoảng cách giữa các nút truyền $\mathcal{X}$ ($\mathcal{X} \in \{\mathcal{S}, \text{IRS}\}$) và nút nhận $\mathcal{Y}$ ($\mathcal{Y} \in \{\text{IRS}, \mathcal{D}\}$). Tham số fading lớn $\lambda_{xy} = PL_0 \left(\frac{d_{xy}}{d_0} \right)^{-\eta}$ với PL_0 đại diện cho suy giảm công suất trung bình tại khoảng cách tham chiếu d_0 ; d_{xy} khoảng cách giữa $\mathcal{X}$ và $\mathcal{Y}$.

Lưu ý rằng, hệ thống được phân tích trong trường gần với kênh truyền fading Rayleigh, nơi fading lớn có thể ảnh hưởng đáng kể, đặc biệt là trong kịch bản xấu nhất. Tuy nhiên, khi có IRS, tác động của fading lớn sẽ giảm nhờ khả năng giảm suy hao tín hiệu (đặc biệt khi có Line-of-Sight - LoS) và cải thiện chất lượng tín hiệu. Trong môi trường gần, nếu có LoS và IRS được bố trí tối ưu, fading nhỏ sẽ ít bị ảnh hưởng hơn nhờ tín hiệu được khuếch đại qua các phản xạ từ IRS [99]. Vì vậy, nghiên cứu mở rộng hệ thống SPC qua kênh truyền LoS sẽ là một trong những hướng nghiên cứu tiếp theo của chúng tôi.

b. Tỷ số tín hiệu trên nhiễu tức thời tại $\mathcal{D}$:

Từ (4.1), SNR tức thời tại $\mathcal{D}$ nhận được như sau:

$$\gamma_D = \frac{P \left| \sum_{i=1}^L \alpha_i \beta_i e^{j(\phi_i - \theta_i - \varphi_i)} \right|^2}{N_0}. \quad (4.4)$$

Để SNR trong (4.4) đạt giá trị tối ưu, IRS điều chỉnh góc dịch chuyển pha để đảm bảo rằng $\phi_i = \theta_i + \varphi_i$ [71]. Bằng cách đặt $\mathcal{A} = \sum_{i=1}^L \alpha_i \beta_i$, phương trình (4.4) được rút gọn lại sau đây:

$$\gamma_D = \mathcal{A}^2 \bar{\gamma}, \quad (4.5)$$

với $\bar{\gamma} = \frac{P}{N_0}$ là SNR của nguồn phát.

Do các hệ số α_i và β_i trong (4.4) là các RVs độc lập và tuân theo phân phối Rayleigh, nên $\mathcal{A}$ là tổng của L RVs độc lập theo phân phối Double Rayleigh.

Khi L đủ lớn (tức $L \gg 1$), theo định lý giới hạn trung tâm, phân phối của $\mathcal{A}$ xấp xỉ phân phối Gauss [100, chương 7]. Ngoài ra, Hình 4.2 chứng tỏ rằng khi L càng lớn ($L \geq 10$) hàm phân phối của $\mathcal{A}$ gần với phân phối Gauss. Do có đặc điểm gần giống phân phối Gauss, phân phối của $\mathcal{A}$ được xấp xỉ chính xác bằng chuỗi Laguerre. Chuỗi Laguerre là một phương pháp hiệu quả giúp xấp xỉ hàm mật độ xác suất phức tạp chính xác hơn. Do đó, hàm PDF của $\mathcal{A}$ có dạng như sau [101, CT. (2.76)] có dạng như sau:

$$f_{\mathcal{A}}(x) = \frac{x^a}{b^{a+1} \Gamma(a+1)} \exp\left(-\frac{x}{b}\right), \quad (4.6)$$

và

$$F_{\mathcal{A}}(x) = \frac{\Upsilon\left(a+1, \frac{x}{b}\right)}{\Gamma(a+1)}, \quad (4.7)$$

với các tham số sau:

$$a = \frac{k_1^2}{k_2} - 1, \quad (4.8)$$

$$b = \frac{k_2}{k_1}, \quad (4.9)$$

với

$$k_1 = \frac{L\pi}{4} \sqrt{\frac{1}{d_1^\eta d_2^\eta}}, \quad (4.10)$$

$$k_2 = \frac{L}{d_1^\eta d_2^\eta} \left(1 - \frac{\pi^2}{16} \right). \quad (4.11)$$

Hàm CDF và PDF của SNR tức thời tại $\mathcal{D}$ và công thức (4.10) và (4.11) được chứng minh sau đây:

Từ (4.5), hàm CDF của của SNR tức thời tại $\mathcal{D}$ được tính như sau:

$$\begin{aligned} F_{\gamma_{\mathcal{D}}}(\gamma) &= \Pr(\gamma_{\mathcal{D}} \leq \gamma) \\ &= \Pr\left(\mathcal{A} \leq \sqrt{\frac{\gamma}{\bar{\gamma}}}\right) = F_{\mathcal{A}}\left(\sqrt{\frac{\gamma}{\bar{\gamma}}}\right). \end{aligned} \quad (4.12)$$

Từ (4.12) và (4.7), hàm CDF của $\gamma_{\mathcal{D}}$ thu được như:

$$\begin{aligned} F_{\gamma}(\gamma) &= \frac{\Upsilon\left(a+1, \frac{1}{b}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right)}{\Gamma(a+1)} \\ &\stackrel{(i)}{=} 1 - \frac{\Gamma\left(a+1, \frac{1}{b}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right)}{\Gamma(a+1)}, \end{aligned} \quad (4.13)$$

với phép rút gọn (i) sử dụng công thức $\gamma(x, y) + \Gamma(x, y) = \Gamma(x)$ trong [85, CT. (8356.3)].

Tiếp tục lấy đạo hàm của (4.13), hàm PDF của $\gamma_{\mathcal{D}}$ tính được như sau:

$$f_{\gamma}(\gamma) = \frac{1}{2b^{a+1}\Gamma(a+1)\bar{\gamma}^{\frac{a+1}{2}}} \gamma^{\left(\frac{a-1}{2}\right)} \exp\left(\frac{1}{b}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right). \quad (4.14)$$

Ngoài ra, hai tham số k_1 và k_2 dựa vào [101, CT. (2.74)], được tính như sau:

$$k_1 = \mathbb{E}[\mathcal{A}]. \quad (4.15)$$

$$k_2 = \mathbb{V}[\mathcal{A}]. \quad (4.16)$$

Bằng cách thế $\mathcal{A}$ vào (4.15) kết quả thu được như sau:

$$\begin{aligned} k_1 &= \mathbb{E}\left[\sum_{i=1}^L \alpha_i \beta_i\right] \\ &\stackrel{(a)}{=} \sum_{i=1}^L \mathbb{E}[\alpha_i] \times \mathbb{E}[\beta_i] \\ &= \sum_{i=1}^L \left(\frac{1}{2}\sqrt{\frac{\pi}{d_1^\eta}}\right) \times \left(\frac{1}{2}\sqrt{\frac{\pi}{d_2^\eta}}\right) \\ &= \frac{\pi L}{4} \sqrt{\frac{1}{d_1^\eta d_2^\eta}}, \end{aligned} \quad (4.17)$$

với (a) biểu thị phép tính trung bình của hai biến α_i và β_i .

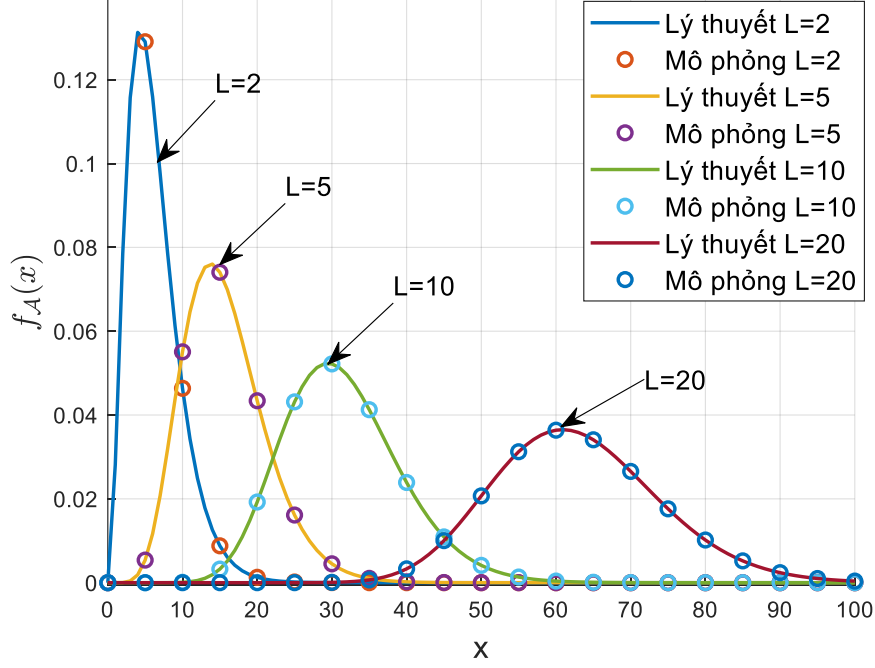
Hơn nữa, ta có $\mathbb{V}[\mathcal{A}] = \mathbb{E}[\mathcal{A}^2] - (\mathbb{E}[\mathcal{A}])^2$ và (4.16), k_2 được tính như sau:

$$\begin{aligned} k_2 &= \mathbb{E}\left[\left(\sum_{i=1}^L \alpha_i \beta_i\right)^2\right] - k_1^2 \\ &= \sum_{i=1}^L \mathbb{E}[|\alpha_i|^2] \times \mathbb{E}[|\beta_i|^2] + \sum_{i=1}^L \mathbb{E}[|\alpha_i|] \times \mathbb{E}[|\beta_i|] \times \sum_{\substack{j=1 \\ j \neq i}}^L \mathbb{E}[|\alpha_j|] \times \mathbb{E}[|\beta_j|] - k_1^2. \end{aligned} \quad (4.18)$$

Bằng cách thay thế $\mathbb{E}[|\alpha_i|^2] = \frac{1}{d_1^\eta}$, $\mathbb{E}[|\beta_i|^2] = \frac{1}{d_2^\eta}$ và (4.17) vào (4.18) (do $|\alpha_i|^2$,

và $|\beta_i|^2$ là các biến ngẫu nhiên phân phối mũ [102]), k_2 thu được như sau:

$$\begin{aligned}
k_2 &= \frac{L}{d_1^\eta d_2^\eta} + L(L-1) \left(\frac{1}{2} \sqrt{\frac{\pi}{d_1^\eta}} \right)^2 \left(\frac{1}{2} \sqrt{\frac{\pi}{d_2^\eta}} \right)^2 - \left(\frac{\pi L}{4\sqrt{d_1^\eta d_2^\eta}} \right)^2 \\
&= \frac{L}{d_1^\eta d_2^\eta} \left(1 - \frac{\pi^2}{16} \right).
\end{aligned} \tag{4.19}$$

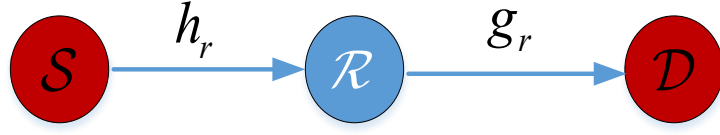


Hình 4.2. Hàm PDF của $\mathcal{A}$ với các giá trị L khác nhau.

4.2.2 Hệ thống chuyển tiếp qua nút chuyển tiếp

Mô hình hệ thống SPC thông qua nút chuyển tiếp ($\mathcal{R}$) được trình bày chi tiết như trong Hình 4.3. Trong hệ thống truyền tin này, chúng ta giả định về kênh truyền giống như mô hình SPC qua IRS. Tuy nhiên, quá trình truyền tin từ $\mathcal{S}$ đến $\mathcal{D}$ được thực hiện trong hai khoảng thời gian liên tiếp được hỗ trợ bởi $\mathcal{R}$.

Ở khoảng thời gian thứ nhất, tín hiệu cần truyền đến $\mathcal{D}$ sẽ được $\mathcal{S}$ truyền tới $\mathcal{R}$. Sau đó, $\mathcal{R}$ xử lý tín hiệu này bằng cách áp dụng kỹ thuật FDF để chuyển tiếp đến $\mathcal{D}$.



Hình 4.3: Mô hình SPC thông qua chuyển tiếp FDF.

SNR toàn trình của hệ thống được biểu diễn bên dưới sau đây:

$$\gamma_{DF} = \min(\gamma_{SR}, \gamma_{RD}), \quad (4.20)$$

với γ_{SR} và γ_{RD} chỉ lần lượt SNR của kênh truyền từ $S \rightarrow R$ và từ $R \rightarrow D$, tương ứng.

Bởi vì tất cả các kênh truyền từ $S \rightarrow R$ và từ $R \rightarrow D$ được giả định là fading Rayleigh hoàn toàn độc lập nhau và có phân phối đồng dạng với nhau (i.i.d.). Do đó, hàm CDF của γ_{DF} được dẫn ra sau đây [66]:

$$\begin{aligned} F_{\gamma_{DF}}(\gamma) &= \Pr(\gamma_{DF} \leq \gamma) \\ &= 1 - \exp\left(-\left(\frac{\gamma}{\bar{\gamma}_{SR}} + \frac{\gamma}{\bar{\gamma}_{RD}}\right)\right), \end{aligned} \quad (4.21)$$

với $\bar{\gamma}_{SR} = \frac{P_S}{N_0} \mathbb{E}[|h_r|^2] = \frac{P_S}{N_0} \Omega_{SR}$ và $\bar{\gamma}_{RD} = \frac{P_R}{N_0} \mathbb{E}[|g_r|^2] = \frac{P_R}{N_0} \Omega_{RD}$ là SNR trung bình của đường truyền từ $S \rightarrow R$ và từ $R \rightarrow D$, tương ứng. Trong đó, P_S và P_R chỉ công suất phát của S và R , tương ứng; N_0 là công suất nhiễu Gauss tại R và D ; h_r , g_r , Ω_{SR} và Ω_{RD} chỉ các hệ số kênh và độ lợi kênh trung bình từ $S \rightarrow R$ và $R \rightarrow D$, tương ứng.

❖ Nếu R sử dụng kỹ thuật AF để truyền tín hiệu đến D , SNR toàn trình trong trường hợp này được biểu diễn như sau:

$$\gamma_{AF} = \frac{\gamma_{SR} \gamma_{RD}}{\gamma_{SR} + \gamma_{RD} + 1}. \quad (4.22)$$

Dựa vào [103, CT(19)], hàm CDF của γ_{AF} được dẫn ra như sau:

$$F_{\gamma_{AF}}(\gamma) = 1 - \frac{2\gamma}{\sqrt{\gamma_{SR}\gamma_{RD}}} \exp\left(-\left(\frac{\gamma}{\gamma_{SR}} + \frac{\gamma}{\gamma_{RD}}\right)\right) \times K_1\left(\frac{2\gamma}{\sqrt{\gamma_{SR}\gamma_{RD}}}\right), \quad (4.23)$$

với $K_1(\cdot)$ là hàm Bessel K bậc 1.

4.3 Phân tích hiệu năng hệ thống

Để đánh giá và so sánh hiệu năng giữa hệ thống SPC qua IRS và hệ thống SPC qua $\mathcal{R}$, chúng ta trình bày từng bước tính BLER dạng chính xác và dạng xấp xỉ chung cho cả hai hệ thống SPC qua IRS và qua $\mathcal{R}$.

Giả sử hệ thống đề xuất cần truyền β bits thông tin qua k kênh sử dụng ($k \geq 100$ CUs) với tốc độ truyền ở chặng thứ $\mathcal{X}$ tương ứng là $\mathcal{R}_{\mathcal{X}} = \frac{\beta_{\mathcal{X}}}{k}$. Dựa vào phân tích

trong (2.9), BLER trung bình ở chặng thứ $\mathcal{X}$ được dẫn ra bên dưới như sau:

$$\bar{\varepsilon}_{e2e} = v\sqrt{k} \int_{\rho_L}^{\rho_H} F_{\gamma_{\mathcal{X}}}(\gamma) d\gamma, \quad (4.24)$$

với $\gamma_{\mathcal{X}}$ là SNR ở chặng thứ $\mathcal{X}$.

Bằng cách thay thế (4.7) vào (4.24), e2e BLER của hệ thống SPC qua IRS thu được sau đây:

$$\bar{\varepsilon}_{e2e}^{IRS} = v\sqrt{k} \left[\rho_H - \rho_L - \frac{2}{\Gamma(a+1)} \left(\Delta(\sqrt{\rho_H}) - \Delta(\sqrt{\rho_L}) \right) \right], \quad (4.25)$$

với

$$\Delta(y) = \frac{1}{2} \left[y^2 \Gamma\left(a+1, \frac{y}{b\sqrt{\gamma}}\right) - \frac{y^{a+2}}{(b\sqrt{\gamma})^a} e^{\frac{-y}{b\sqrt{\gamma}}} - \left(\frac{1}{b\sqrt{\gamma}}\right)^2 \Gamma\left(a+2, \frac{y}{b\sqrt{\gamma}}\right) \right]. \quad (4.26)$$

(Công thức (4.25) được chứng minh trong Phụ lục D).

Tương tự, bằng cách thế (4.21) vào (4.24) và sau đó thực hiện lấy tích phân của (4.24), e2e BLER của hệ thống SPC qua $\mathcal{R}$ thu được như sau:

$$\bar{\varepsilon}_{e2e}^{FDF} = v\sqrt{k} \left[(\rho_H - \rho_L) + \left(\frac{\bar{\gamma}_{SR} \cdot \bar{\gamma}_{RD}}{\bar{\gamma}_{SR} + \bar{\gamma}_{RD}} \right) \times \exp \left(- \left(\frac{\bar{\gamma}_{SR} + \bar{\gamma}_{RD}}{\bar{\gamma}_{SR} \bar{\gamma}_{RD}} \right) \times (\rho_H - \rho_L) \right) \right]. \quad (4.27)$$

Tiếp theo, bằng cách áp dụng công thức trong [85, CT. (8354.1)] cho (4.7), kết quả hàm CDF xấp xỉ của γ được rút gọn là:

$$F_\gamma(\gamma) \simeq \frac{1}{\Gamma(a+1)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{b} \sqrt{\frac{\gamma}{\bar{\gamma}}} \right)^{a+1+m}}{m!(a+1+m)}, \quad (4.28)$$

với m là số bậc khai triển.

Thay thế (4.28) vào (4.24) và khai triển công thức xấp xỉ $1 - \exp(-x) \sim x$ khi $x \rightarrow 0$ cho (4.26), kết quả BLER dạng xấp xỉ cho hai trường hợp SPC lần lượt qua IRS và qua $\mathcal{R}$ sau đây:

$$\bar{\varepsilon}_{e2e}^{IRS} \simeq \frac{v\sqrt{k}}{\Gamma(a+1)} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(a+1+m)(a+2+m)} \times \left(\left(\frac{1}{b} \sqrt{\frac{\rho_H}{\bar{\gamma}}} \right)^{a+2+m} - \left(\frac{1}{b} \sqrt{\frac{\rho_L}{\bar{\gamma}}} \right)^{a+2+m} \right). \quad (4.29)$$

và

$$\bar{\varepsilon}_{e2e}^{FDF} \simeq v\sqrt{k} \left(\frac{\bar{\gamma}_{SR} + \bar{\gamma}_{RD}}{\bar{\gamma}_{SR} \bar{\gamma}_{RD}} \right) (\rho_H - \rho_L). \quad (4.30)$$

Ngoài ra, dựa vào xấp xỉ tích phân Riemann, e2e BLER cho chuyển tiếp qua $\mathcal{R}$ sử dụng kỹ thuật AF được tính như sau:

$$\bar{\varepsilon}_{e2e}^{AF} \simeq F_{\gamma_{AF}}(\theta) = 1 - \frac{2\theta}{\sqrt{\bar{\gamma}_{SR}\bar{\gamma}_{RD}}} \exp\left(-\left(\frac{\theta}{\bar{\gamma}_{SR}} + \frac{\theta}{\bar{\gamma}_{RD}}\right)\right) \times K_1\left(\frac{2\theta}{\sqrt{\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}\right). \quad (4.30)$$

4.4 Mô phỏng và thảo luận

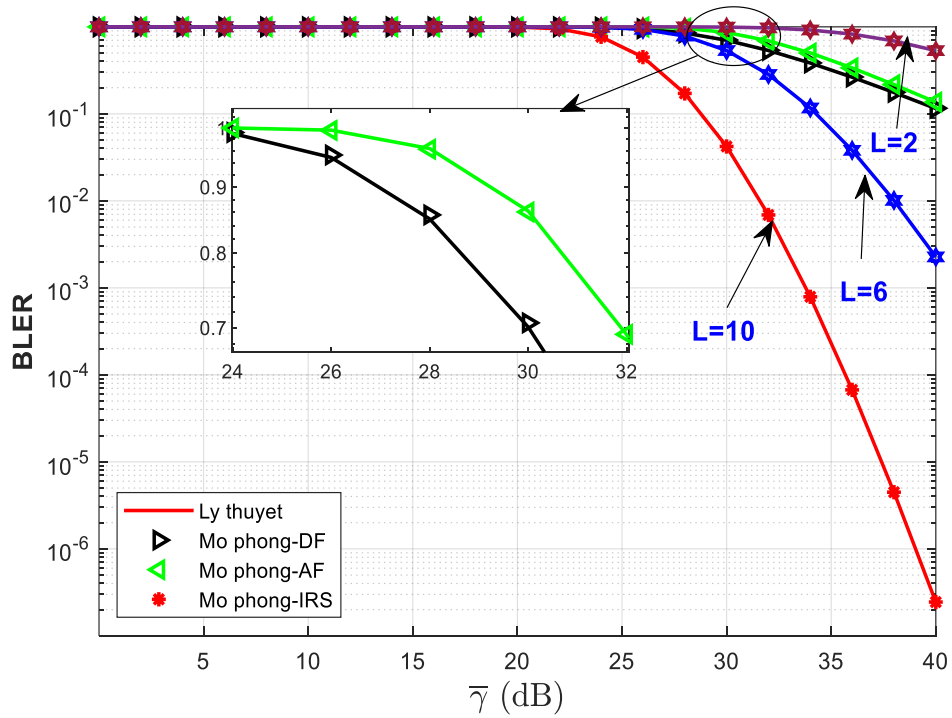
Trong phần này, chúng ta tiến hành mô phỏng Monte Carlo sử dụng phần mềm Matlab [104], nhằm xác minh lại những kết luận lý thuyết đã được dẫn ra ở mục 4.2. Để đảm bảo sự công bằng, các thông số mô phỏng được thiết lập giống nhau cho cả hai hệ thống SPC sử dụng IRS và $\mathcal{R}$, bao gồm số bit thông tin $\beta = 256$ bits; độ dài khối tin $k = 200$ CUs; khoảng cách truyền chuẩn hóa giữa nguồn phát $\mathcal{S}$ và nguồn nhận $\mathcal{D}$ là $D = 10$; đồng thời phân bố khoảng cách của các kênh truyền từ $\mathcal{S} \rightarrow \text{IRS} / \mathcal{R}$ và $\text{IRS} / \mathcal{R} \rightarrow \mathcal{D}$ là bằng nhau và ký hiệu d_1 và d_2 . Đối với hệ thống sử dụng $\mathcal{R}$ xem xét cả hai chế độ FDF và AF. Ngoài ra, công suất của nguồn phát được phân bố đồng đều cho cả hai hệ thống SPC sử dụng IRS và $\mathcal{R}$. Để mô phỏng sự suy giảm đường truyền, một mô hình suy hao đơn giản được áp dụng cho cả hai hệ thống với hệ số suy giảm đường truyền $\eta = 3$ [93], $PL_0 = -30\text{dB}$ [35, 105] là suy hao công suất tại khoảng cách tham chiếu $d_0 = 1\text{ m}$, và $d = \{d_1, d_2\}$ là tập hợp các khoảng cách của kênh truyền.

Đầu tiên, BLER trong các mô hình SPC qua IRS và $\mathcal{R}$ được biểu diễn như Hình 4.4. Trong đó, BLER được thể hiện như một hàm của $\bar{\gamma}$, với sự biến đổi yếu tố phản xạ L của IRS tại các giá trị $L = 2$, $L = 6$ và $L = 10$.

Quan sát Hình 4.4, kết quả mô phỏng và lý thuyết cho thấy sự tương đồng rõ rệt. Cụ thể, ba đường, tức là đường đại diện cho mô phỏng (các dấu chấm) và đường lý thuyết (đường nét liền) đều khớp với nhau. Hơn nữa, trong hệ thống SPC qua IRS, BLER giảm khi giá trị $\bar{\gamma}$ tăng. Khi $\bar{\gamma}$ ở giá trị cao, tốc độ giảm của BLER trở nên nhanh hơn, phù hợp với công thức lý thuyết (4.25), khi đó hàm $\bar{\varepsilon}_{e2e}^{\text{IRS}}$ giảm mạnh theo $\bar{\gamma}$. Tương tự, đối với mô hình SPC thông qua nút $\mathcal{R}$, BLER giảm dần khi $\bar{\gamma}$ tăng, nhưng tốc độ giảm không mạnh như hệ thống sử dụng IRS.

Điều này minh chứng rằng việc tích hợp IRS giúp cải thiện đáng kể hiệu năng truyền thông, đặc biệt khi L tăng từ 6 đến 10. Hình 4.4 minh họa rằng BLER của hệ thống qua IRS cải thiện đáng kể hơn (giảm nhanh hơn) so với hệ thống FDF và AF, khi xét $L = 6$ và $L = 10$. Đặc biệt, khi tăng giá trị của L thì BLER càng giảm, tức là hiệu năng cải thiện.

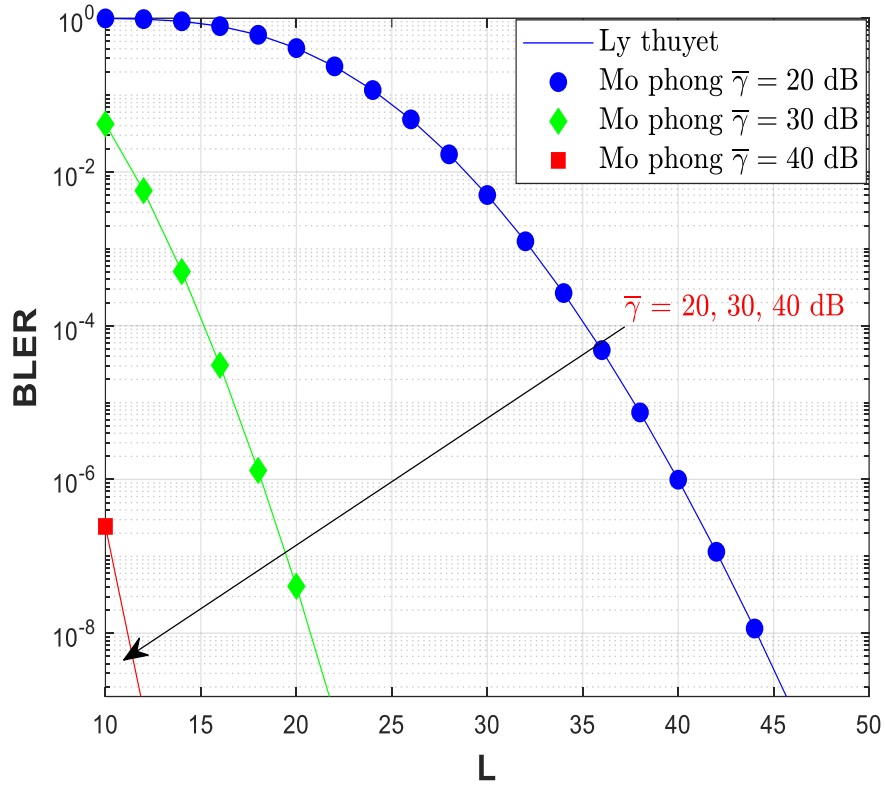
Tiếp theo, khảo sát sự tác động của BLER vào yếu tố phản xạ L của hệ thống SPC qua IRS ứng với ba trường hợp $\bar{\gamma}$ tương ứng như $\bar{\gamma} = 20\text{ dB}$, $\bar{\gamma} = 30\text{ dB}$ và $\bar{\gamma} = 40\text{ dB}$ như trong Hình 4.5.



Hình 4.4: BLER giữa hai mô hình SPC qua IRS và qua $\mathcal{R}$ với $L = 2$, $L = 6$ và $L = 10$, tương ứng.

Quan sát Hình 4.5 chúng ta rõ ràng nhận thấy rằng sự tăng của cả hai tham số L và $\bar{\gamma}$, BLER giảm đáng kể. Tuy nhiên, khi $\bar{\gamma}$ tăng, công suất của nguồn phát cũng tăng theo, dẫn đến can nhiễu cũng sẽ tăng. Vì vậy, để cải thiện hiệu năng của hệ thống với nguồn công suất giữ nguyên, có thể xem xét việc tăng các phần tử phản xạ L của IRS để đạt được mức BLER như mong muốn.

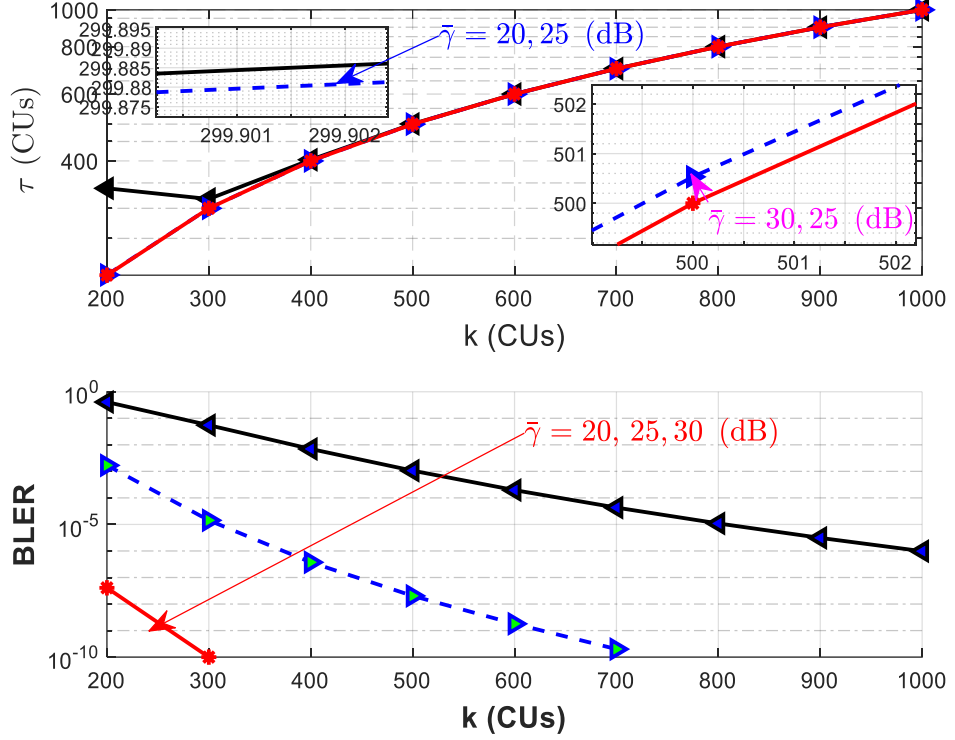
Cuối cùng, chúng ta sẽ phân tích sự tác động của độ dài gói tin k lên BLER và độ trễ (τ) của hệ thống SPC sử dụng IRS với các giá trị tăng của $\bar{\gamma}$ lần lượt như 20dB , 25 dB , 30 dB , như trong Hình 4.6. Độ trễ τ được định nghĩa là thời gian nguồn phát truyền thành công tín hiệu đến nguồn nhận $\frac{k}{1 - \bar{\epsilon}_{e2e}^{\text{IRS}}}$.



Hình 4.5: BLER theo số yếu tố phản xạ L trong hệ thống SPC qua IRS.

Quan sát Hình 4.6 ta thấy rõ một sự đánh đổi giữa độ tin cậy và độ trễ trong hệ thống SPC sử dụng IRS. Cụ thể, k và $\bar{\gamma}$ đều có tác động tích cực đến độ tin cậy, thể hiện qua việc giảm mạnh BLER. Tuy nhiên, khi k tăng độ trễ τ cũng tăng tỉ lệ thuận với độ dài gói tin, dẫn đến thời gian truyền tin lớn hơn. Ở mức $\bar{\gamma} = 30$ dB, BLER giảm xuống giá trị cực thấp 10^{-10} khi $k \geq 300$. Điều này chứng minh rằng hệ thống đạt được độ tin cậy rất cao khi công suất tín hiệu lớn và gói tin dài. Tuy

nhiên, việc độ trễ tăng theo k đòi hỏi phải cân nhắc kỹ lưỡng trong thiết kế để đáp ứng yêu cầu cụ thể của các ứng dụng.



Hình 4.6: Sự ảnh hưởng BLER và độ trễ vào độ dài gói tin k trong hệ thống SPC qua IRS.

4.5 Kết luận Chương 4

Chương 4 đã trình bày những điểm mới trong Luận án như sau:

- Điểm thứ nhất là thiết kế mặt IRS hỗ trợ cho hệ thống truyền SPC trong mạng chuyên tiếp. Từ đó, chúng ta cũng đưa ra phương pháp phân tích hiệu năng mới thông qua các tham số như BLER dạng tường minh và dạng xấp xỉ ở vùng SNR cao.
- Điểm thứ hai là đề xuất so sánh hiệu năng của mô hình SPC thông qua IRS với mô hình SPC qua $\mathcal{R}$. Chúng ta đã thực hiện so sánh thông qua độ lợi phân tập của hệ thống. Kết quả chứng minh rằng hiệu năng của hệ thống SPC được hỗ trợ của IRS cải thiện đáng kể và đạt độ phân tập cao hơn hệ thống qua $\mathcal{R}$.

- Cuối cùng, hiệu năng của hệ thống SPC được hỗ trợ bởi IRS cải thiện hơn hệ thống DF và AF qua $\mathcal{R}$, khi công suất phát được phân bổ như nhau.

Mô hình trong Chương 4 mang lại hiệu quả cao trong mạng 5G và các mạng tương lai, đặc biệt là trong những môi trường truyền tín hiệu khó khăn, như: 1) Các khu vực có địa hình phức tạp, chẳng hạn như đô thị và vùng đồi núi; 2) Các tình huống với nguồn năng lượng phát sóng hạn chế và phạm vi phủ sóng nhỏ. Tuy nhiên, mô hình này cũng đối mặt với một số thách thức khi triển khai thực tế, bao gồm chi phí lắp đặt và bảo trì cao, cùng với yêu cầu khắt khe về độ chính xác trong hệ thống quản lý và điều khiển IRS.

Nội dung của Chương 4 được tóm lược trong hai bài báo, đã được đăng trên tạp chí Khoa học và Công nghệ Thông tin và Truyền thông (JSTIC), chi tiết như sau:

1. **Nguyễn Thị Yến Linh**, V. N. Q. Bao, Phạm Ngọc Sơn, "*Phân Tích Độ Lợi Phân Tập Cho Mạng Chuyển Tiếp Qua Mặt Phản Xạ Thông Minh Và Nút Chuyển Tiếp Trong Truyền Thông Gói Tin Ngắn*," Journal of Science and Technology on Information and Communications (JSTIC), vol. 01, no. 04B, pp. 56-64, Dec. 2021.
2. **Nguyễn Thị Yến Linh**, V. N. Q. Bao, Phạm Ngọc Sơn, "*Dự Đoán Công Suất Tối Ưu Nút nguồn Phát Của Hệ Thống Truyền Gói Tin Ngắn Qua Mặt Phản Xạ Thông Minh Bằng Kỹ Thuật Học Sâu*," Journal of Science and Technology on Information and Communications (JSTIC), vol. 1, no. 1A, pp. 11-18, Apr. 2023.

Chương 5. THIẾT KẾ HỆ THỐNG TRUYỀN THÔNG TIN GÓI NGẮN TRONG MẠNG ĐA TRUY NHẬP PHI TRỰC GIAO SỬ DỤNG MẶT PHẢN XẠ THÔNG MINH

5.1 Giới thiệu

Mục tiêu chính của Chương 5 là thực hiện mục tiêu thứ hai của Luận án, đó là giải quyết các vấn đề thách thức hiện nay trong nghiên cứu SPC, đáp ứng đồng thời hai yêu cầu quan trọng: độ tin cậy cao và độ trễ cực thấp cho các hệ thống không dây thế hệ mới. Mặc dù mô hình trong Chương 4 đã chỉ ra sự cải thiện hiệu năng khi thay thế nút chuyển tiếp bằng tấm IRS, nhưng với yêu cầu về thời gian thực thi ngắn trong các hệ thống đa kết nối người dùng (đặc biệt là IoT và mạng công nghiệp), cũng như hiệu quả năng lượng, mô hình trong Chương 4 vẫn gặp một số hạn chế đáng kể.

Hơn nữa, trong thời gian gần đây, học sâu đã được chứng minh là một phương pháp hiệu quả để giải quyết vấn đề về thời gian thực thi ngắn và giảm thiểu tính toán trong các nghiên cứu về SPC [45, 46, 106]. Cụ thể, nghiên cứu trong công trình [45] đã áp dụng học sâu để dự đoán BLER trong mạng NOMA kết hợp với vô tuyến nhận thức. Kết quả cho thấy mô hình học sâu mang lại sai số dự đoán thấp nhất và thời gian dự đoán ngắn nhất so với các mô hình hồi quy đa đầu ra khác. Các nghiên cứu tiếp theo cũng đã chứng minh sự cải tiến về thời gian dự đoán và độ chính xác, điển hình như công trình [106] áp dụng học sâu để dự đoán BLER, tổng thông lượng và hiệu quả năng lượng trong mạng IoT với vô tuyến nhận thức. Một công trình khác [46], cũng đã sử dụng học sâu để mở rộng khả năng dự đoán thông lượng cho hệ thống SPC trong mạng IoT nhiều chặng. Hầu hết các công trình trên cho thấy hiệu quả của học sâu trong dự đoán hiệu năng (qua các tham số như BLER, thông lượng, độ trễ).

Dựa trên các kết quả này, trong Chương 5 của Luận án, ngoài các cải tiến thiết kế như sử dụng đa an-ten tạo chùm tại nguồn phát, kết hợp với NOMA và sử dụng IRS, hệ thống đề xuất phương pháp học sâu để giải quyết các vấn đề sau: 1) dự

đoán hiệu năng hệ thống (dự đoán BLER); 2) dự đoán công suất phân bổ nguồn phát. Cuối cùng Chương 5 được cấu trúc như sau: đầu tiên Mục 5.2 giới thiệu mô hình SPC mới cùng với đề xuất mô hình tính toán cho hệ thống. Tiếp theo, Mục 5.3 thực hiện phân tích và xác định các tham số như BLER, thông lượng và độ trễ trên kênh fading Rayleigh. Mục 5.4 tập trung vào việc phát triển mạng nơ ron sâu và đề xuất phương pháp dự đoán BLER và công suất phân bổ nguồn phát cho hệ thống. Cuối cùng, Mục 5.5 chứng minh lợi ích mạng DNN trong việc dự đoán hiệu năng và công suất hệ thống, qua mô phỏng Monte Carlo.

5.2 Mô hình hệ thống

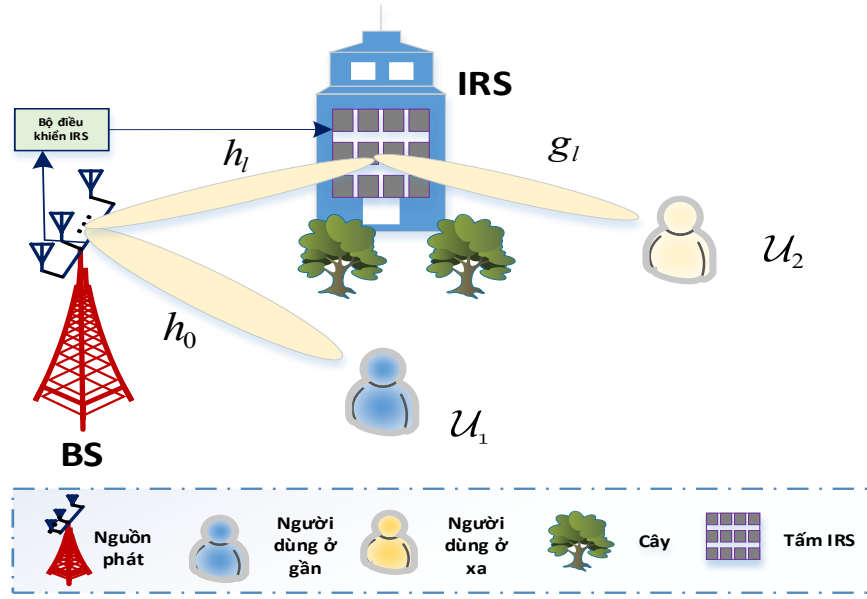
Hình 5.1. là mô hình hệ thống mạng NOMA được hỗ trợ bởi IRS trong SPC. Cụ thể, một nguồn phát (BS) gồm một mảng N an-ten, sử dụng kỹ thuật NOMA để truyền tín hiệu đồng thời với hai người dùng đơn an-ten như ở gần $\mathcal{U}_1$ và ở xa $\mathcal{U}_2$. Một tấm IRS bao gồm L phần tử phản xạ thụ động, được sử dụng để cải thiện chất lượng truyền tín hiệu từ nguồn BS đến người dùng ở xa $\mathcal{U}_2$. Để tối ưu tín hiệu truyền sóng đến $\mathcal{U}_1$, BS sử dụng kỹ thuật định dạng chùm tia, tức BS có thể điều chỉnh mảng an-ten sao cho chùm bức xạ chính nằm dọc theo trục của mảng.

Cụ thể, mảng end-fire với góc bức xạ cực đại 0° hoặc 180° được giả định trong mô hình đề xuất [107]. Hơn nữa, đường truyền trực tiếp giữa BS và $\mathcal{U}_2$ không tồn tại do sự cản trở của vật cản trong môi trường như cây cối và tòa nhà. Ngoài ra, mọi kênh truyền trong hệ thống đều là kênh fading Rayleigh không phụ thuộc và được phân bố tương tự nhau, với cùng hệ số suy hao tín hiệu.

Theo kỹ thuật NOMA, BS sẽ phát tín hiệu tổng hợp $x_{BS} = \sum_{i=1}^2 \sqrt{\alpha_i P_{BS}} x_i$ đến $\mathcal{U}_i$

với $i \in \{1, 2\}$, trong đó x_i là dữ liệu mà nguồn BS muốn gửi đến $\mathcal{U}_i$, $\mathbb{E}\{|x_i|^2\} = 1$;

P_{BS} là công suất phát của BS và α_i là hệ số phân chia công suất cho người dùng $\mathcal{U}_i$, tương ứng.



Hình 5.1: Mô hình hệ thống mạng NOMA-IRS.

Do chùm bức xạ chính từ mảng end-fire ở nguồn BS dọc theo phương truyền tới $\mathcal{U}_1$ nên công suất phân bổ theo thứ tự ở $\mathcal{U}_1$ và $\mathcal{U}_2$ như $\alpha_1 > \alpha_2 \geq 0$ với $\alpha_1 + \alpha_2 = 1$. Do đó, tín hiệu thu được tại $\mathcal{U}_1$ và $\mathcal{U}_2$ có dạng như sau:

$$y_{\mathcal{U}_1} = \left| AF(\theta_{s\mathcal{U}_1}) \right| h_0 x_{BS} + n_1, \quad (5.1)$$

và

$$y_{\mathcal{U}_2} = \left| AF(\theta_{s\mathcal{U}_2}) \right| \left(\sum_{\ell=1}^L h_{\ell} r_{\ell} g_{\ell} \right) x_{BS} + n_2, \quad (5.2)$$

với $AF(\theta) = \frac{1}{N} \left\{ \frac{\sin \left[\frac{N}{2} (Kd \cos \theta + b) \right]}{\sin \left[\frac{1}{2} (Kd \cos \theta + b) \right]} \right\}$ là hệ số sắp xếp của mảng N an-ten

[107, CT (6.27)]; $\theta \in \{\theta_{s\mathcal{U}_1}, \theta_{s\mathcal{U}_2}\}$ là các góc bức xạ; b độ lệch kích thích pha giữa

các yếu tố của mảng; $K = \frac{2\pi}{\lambda}$ là hệ số pha; $d = \frac{\lambda(N-1)}{2N}$ là khoảng cách giữa

các an-ten trong mảng an ten; λ là bước sóng của tín hiệu tới; n_i là nhiễu cộng tại $\mathcal{U}_i$ (n_i có phân phối Gauss với giá trị trung bình 0 và phương sai N_0 với

$n_i \sim CN(0, N_0)$). Hơn nữa, hệ số phản xạ của yếu tố phản xạ thứ ℓ là $r_\ell \triangleq |r_\ell| e^{j\psi_\ell}$, ($\ell = 1, 2, \dots, L$) với ψ_ℓ là pha có thể điều chỉnh, $\psi_\ell \in [0, 2\pi)$ và biên độ $|r_\ell| = 1$ [27]; $h_l = d_1^{-\eta/2} A_l e^{-j\phi_l}$ và $g_l = d_2^{-\eta/2} B_l e^{-j\phi_l}$ là các hệ số kênh của các đường truyền BS-IRS và IRS- $\mathcal{U}_2$, với d_1 và d_2 là các khoảng cách giữa BS đến IRS và IRS đến $\mathcal{U}_2$; η là hệ số suy hao đường truyền; A_l , B_l , θ_l và ϕ_l chỉ các biên độ và pha ứng với các hệ số kênh h_l và g_l .

Bằng cách sử dụng phương pháp SIC, $\mathcal{U}_1$ có chất lượng kênh tốt nên sẽ ưu tiên giải mã trực tiếp tín hiệu của nó x_1 , đồng thời xem xét tín hiệu với điều kiện kênh kém như là nhiễu x_2 . Như vậy, tỉ lệ tín hiệu cộng nhiễu (SINR) tại $\mathcal{U}_1$ được tính như:

$$\gamma_{\mathcal{U}_1}^{x_1} = \frac{\alpha_1 \Omega_0 G_1 \bar{\gamma}}{G_1 \Omega_0 \alpha_2 \bar{\gamma} + 1}, \quad (5.3)$$

với $G_1 = |AF(\theta_{s\mathcal{U}_1})|^2$, $\bar{\gamma} = \frac{P}{N_0}$ là SNR phát và $\Omega_0 = |h_0|^2$ là độ lợi kênh ở $\mathcal{U}_1$.

Tại $\mathcal{U}_2$, x_1 sẽ được giải mã đầu tiên, trong đó x_2 được xem xét như là nhiễu. Sau đó, tín hiệu x_1 được trừ đi trước khi $\mathcal{U}_2$ tiến hành quá trình giải mã x_2 . SINR tại $\mathcal{U}_2$ cho quá trình giải mã x_1 và x_2 được diễn đạt tương ứng như sau:

$$\gamma_{\mathcal{U}_2}^{x_1} = \frac{\alpha_1 P |AF(\theta_{s\mathcal{U}_2})|^2 \left(\sum_{l=1}^L h_l g_l r_l \right)^2}{|AF(\theta_{s\mathcal{U}_2})|^2 \left(\sum_{l=1}^L h_l g_l r_l \right)^2 \alpha_2 P + N_0}. \quad (5.4)$$

và

$$\gamma_{\mathcal{U}_2}^{x_2} = \frac{\alpha_2 P |AF(\theta_{s\mathcal{U}_2})|^2 \left(\sum_{l=1}^L h_l g_l r_l \right)^2}{N_0}. \quad (5.5)$$

Chúng ta đặt giả thuyết rằng CSI là hoàn hảo. Do đó, IRS có thể điều khiển sự dịch pha của từng phần tử phản xạ trong IRS sao cho SINR đạt cực đại tại $\mathcal{U}_2$, tức $\psi_l = -(\phi_l + \varphi_l)$. Hơn nữa, chúng ta xem xét vùng trường gần của mảng an-ten trong đó khoảng cách từ BS/ $\mathcal{U}_2$ đến chính giữa IRS nhỏ hơn λ (độ dài sóng của tín hiệu) và kích thước tối đa của IRS được ký hiệu là D .

Bằng cách đặt $G_2 = \left| AF(\theta_{s\mathcal{U}_2}) \right|^2$ và $\Theta = \sum_{l=1}^L h_l g_l r_l = \sum_{l=1}^L d_1^{-\eta/2} d_2^{-\eta/2} A_l B_l$, SINR ở $\mathcal{U}_2$ được trình bày gọn lại như sau:

$$\gamma_{\mathcal{U}_2}^{x_1} = \frac{\alpha_1 G_2 \Theta^2 \bar{\gamma}}{G_2 \alpha_2 \Theta^2 \bar{\gamma} + 1}, \quad (5.6)$$

và

$$\gamma_{\mathcal{U}_2}^{x_2} = \alpha_2 G_2 \Theta^2 \bar{\gamma}. \quad (5.7)$$

5.3 Phân tích hiệu năng hệ thống

5.3.1 Hàm CDF của SINR tại $\mathcal{U}_1$ và $\mathcal{U}_2$

Trong phần này, chúng ta sẽ chi tiết hóa các kiến thức cơ bản cần thiết để tính toán hàm CDF của SINR tại $\mathcal{U}_1$ và $\mathcal{U}_2$. Điều này dựa trên giả thuyết rằng kênh truyền giữa người dùng và nguồn BS là fading Rayleigh. Hệ số kênh Ω_0 là RV, và có phân phối mũ. Do đó, hàm CDF của Ω_0 được mô tả trong [108, CT(6.68)]:

$$F_{\Omega_0}(x) = 1 - \exp(-\lambda_0 x), \quad (5.8)$$

với λ_0 là một tham số của Ω_0 và được mô hình đơn giản như $\lambda_0 = d_0^\eta$, với hệ số suy giảm kênh truyền η và d_0 khoảng cách giữa BS và $\mathcal{U}_1$, tương ứng.

Hơn nữa, dựa vào [108, CT(4.1)] hàm CDF của $\gamma_{\mathcal{U}_1}^{x_1}$ thu được như sau:

$$F_{\gamma_{\mathcal{U}_1}^{x_1}}(x) = \Pr\left(\frac{\alpha_1 \Omega_0 G_1 \bar{\gamma}}{G_1 \Omega_0 \alpha_2 \bar{\gamma} + 1} < x\right) = \begin{cases} 1, & x \geq \frac{\alpha_1}{\alpha_2}, \\ F_{\Omega_0}\left(\frac{x}{G_1 \bar{\gamma}(\alpha_1 - \alpha_2 x)}\right), & x < \frac{\alpha_1}{\alpha_2}. \end{cases} \quad (5.9)$$

Từ (5.9) và (5.8), hàm CDF của $\gamma_{\mathcal{U}_1}^{x_1}$ thu được như sau:

$$F_{\gamma_{\mathcal{U}_1}^{x_1}}(x) = \begin{cases} 1, & x \geq \frac{\alpha_1}{\alpha_2}, \\ 1 - \exp\left(-\lambda_0 \frac{x}{G_1 \bar{\gamma}(\alpha_1 - \alpha_2 x)}\right), & x < \frac{\alpha_1}{\alpha_2}. \end{cases} \quad (5.10)$$

Mệnh đề 1. Hàm CDF của $\gamma_{\mathcal{U}_2}^{x_1}$ và $\gamma_{\mathcal{U}_2}^{x_2}$ ở $\mathcal{U}_2$ được diễn tả như sau:

$$F_{\gamma_{\mathcal{U}_2}^{x_1}}(x) = \begin{cases} 1, & x \geq \frac{\alpha_1}{\alpha_2}, \\ 1 - \frac{\Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{G_2 \bar{\gamma}(\alpha_1 - \alpha_2 x)}}\right)}{\Gamma(a+1)}, & x < \frac{\alpha_1}{\alpha_2}. \end{cases} \quad (5.11)$$

và

$$F_{\gamma_{\mathcal{U}_2}^{x_2}}(x) = 1 - \frac{\Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{\alpha_2 G_2 \bar{\gamma}}}\right)}{\Gamma(a+1)}, \quad (5.12)$$

Trong đó $a = k_1^2 / k_2 - 1$, $b = k_2 / k_1$, $k_1 = \frac{L\pi}{4} \sqrt{1/d_1^\eta d_2^\eta}$ và $k_2 = L / d_1^\eta d_2^\eta \left(1 - \frac{\pi^2}{16}\right)$.

Chứng minh Mệnh đề 1:

Bằng cách xem xét $\Theta = \sum_{l=1}^L h_l g_l r_l = \sum_{l=1}^L d_1^{-\eta/2} d_2^{-\eta/2} A_l B_l$, chúng ta có thể nhận thấy được rằng A_l và B_l là những RV có phân phối Rayleigh. Do đó, tổng Θ của hai RV phân phối Rayleigh liên tục với $L \geq 1$. Hơn nữa, theo định lý giới hạn trung tâm, Θ hội tụ về biến ngẫu nhiên phân phối Gaussian, tức là $\Theta \sim \mathcal{N}\left(L\frac{\pi}{4}, L\left(1 - \frac{\pi^2}{16}\right)\right)$ [30]. Dựa vào [101, CT. (2.76)], hàm PDF của Θ có thể được xấp xỉ bằng phương pháp chuỗi Laguerre như sau:

$$f_{\Theta}(x) = \frac{x^a}{b^{a+1}\Gamma(a+1)} \exp\left(-\frac{x}{b}\right), \quad (5.13)$$

với $a = k_1^2 / k_2 - 1$ và $b = k_2 / k_1$. Hệ số k_1 và k_2 được chứng minh như trong (4.17) và (4.18).

Hơn nữa, bằng cách lấy tích phân (5.13), chúng ta thu được hàm CDF của Θ sau đây:

$$F_{\Theta}(x) = \frac{\gamma\left(a+1, \frac{x}{b}\right)}{\Gamma(a+1)} \stackrel{(i)}{=} 1 - \frac{\Gamma\left(a+1, \frac{x}{b}\right)}{\Gamma(a+1)}, \quad (5.14)$$

với $\gamma(\cdot)$ và $\Gamma(\cdot)$ được định nghĩa trong danh mục các hàm toán học, (i) được thực hiện bằng cách sử dụng [85, CT. (8356-3)].

Từ (5.6) và (5.7) Chúng ta thấy được $\Theta \propto \sqrt{\frac{\gamma_{\mathcal{U}_2}^{x_i}}{\gamma}}$ và bằng cách kết hợp (5.10) và (5.14), hàm CDF của $\gamma_{\mathcal{U}_2}^{x_i}$ và ở $\mathcal{U}_2$ thu được như trong (5.11) và (5.12).

5.3.2 Tỷ lệ lỗi khối

Ở phần này, biểu thức dạng đóng của BLER tại $\mathcal{U}_1$ và $\mathcal{U}_2$ được dẫn ra cho hệ thống đề xuất dưới điều kiện giao tiếp gói tin ngắn. Nguồn BS được giả định rằng truyền một gói tin chứa β bits thông tin thông qua kênh fading bán tĩnh phẳng

[84]. Dựa vào (2.9), BLER trung bình cho quá trình giải mã tín hiệu $\mathcal{X}$ với $\mathcal{X} \in \{x_1, x_2\}$ tại $\mathcal{U}_i$ được viết gọn lại bên dưới như sau:

$$\bar{\varepsilon}_{\mathcal{U}_i}^{\mathcal{X}} \approx v_i \sqrt{k} \int_{\rho_{1L}}^{\rho_{1H}} F_{\gamma_{\mathcal{U}_i}^{\mathcal{X}}}(\gamma) d\gamma, \quad (5.15)$$

Bổ đề 1. Trong trường hợp kênh truyền fading Rayleigh, BLER trung bình ở $\mathcal{U}_1$ có thể đạt được như sau:

$$\bar{\varepsilon}_{\mathcal{U}_1} = \begin{cases} v_1 \sqrt{k} (\rho_{1H} - \rho_{1L}), & \frac{\alpha_1}{\alpha_2} \leq \rho_{1L}, \\ v_1 \sqrt{k} (\rho_{1H} - \rho_{1L} - \mathcal{I}_1), & \rho_{1L} \leq \frac{\alpha_1}{\alpha_2} < \rho_{1H}, \\ v_1 \sqrt{k} (\rho_{1H} - \rho_{1L} - \mathcal{I}_2), & \frac{\alpha_1}{\alpha_2} \geq \rho_{1H}, \end{cases} \quad (5.16)$$

với

$$\mathcal{I}_1 = \frac{\exp\left(\frac{\lambda_0}{\alpha_2 G_1 \bar{\gamma}}\right)}{\alpha_2} \left[(\alpha_1 - \alpha_2 \rho_{1L}) \exp\left(\frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma}}\right) - \frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma}} \Gamma\left(0, \frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{1L})}\right) \right]. \quad (5.17)$$

Và

$$\begin{aligned} \mathcal{I}_2 = & \frac{\exp\left(\frac{\lambda_0}{\alpha_2 G_1 \bar{\gamma}}\right)}{\alpha_2} \left\{ (\alpha_1 - \alpha_2 \rho_{1L}) \exp\left(\frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{1L})}\right) \right. \\ & - (\alpha_1 - \alpha_2 \rho_{1H}) \exp\left(\frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{1H})}\right) \\ & \left. + \frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma}} \left[\Gamma\left(0, \frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{1L})}\right) - \Gamma\left(0, \frac{\lambda_0 \alpha_1}{\alpha_2 G_1 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{1H})}\right) \right] \right\}. \end{aligned}$$

(5.18)

(Bổ đề 1 được chứng minh trong Phụ lục E).

Tiếp theo, BLER trung bình tổng hợp tại $\mathcal{U}_2$ được trình bày qua **Mệnh đề 2** bên dưới như sau:

Mệnh đề 2. Từ **Bổ đề 1** và (5.15), BLER trung bình cho quá trình giải mã tín hiệu x_1 và x_2 tại $\mathcal{U}_2$ nhận được như:

$$\bar{\varepsilon}_{\mathcal{U}_2} \approx \bar{\varepsilon}_{\mathcal{U}_2}^{x_1} + \bar{\varepsilon}_{\mathcal{U}_2}^{x_2} - \min \left\{ \bar{\varepsilon}_{\mathcal{U}_2}^{x_1}, \bar{\varepsilon}_{\mathcal{U}_2}^{x_2} \right\}, \quad (5.19)$$

với $\bar{\varepsilon}_{\mathcal{U}_2}^{x_1}$ và $\bar{\varepsilon}_{\mathcal{U}_2}^{x_2}$ là BLER trung bình cho quá trình giải mã tín hiệu x_1 và x_2 , tương ứng tại $\mathcal{U}_2$.

Chứng minh Mệnh đề 2 như sau:

Sau khi thực hiện thành công SIC để giải mã x_1 và x_2 , chúng ta có thể tính BLER trung bình tổng hợp tại $\mathcal{U}_2$ như sau:

$$\begin{aligned} \bar{\varepsilon}_{\mathcal{U}_2} &= \mathbb{E} \left\{ \varepsilon_{\mathcal{U}_2}^{x_1} + (1 - \varepsilon_{\mathcal{U}_2}^{x_1}) \varepsilon_{\mathcal{U}_2}^{x_2} \right\} \\ &= \bar{\varepsilon}_{\mathcal{U}_2}^{x_1} + \bar{\varepsilon}_{\mathcal{U}_2}^{x_2} - \mathbb{E} \left\{ \varepsilon_{\mathcal{U}_2}^{x_1} \varepsilon_{\mathcal{U}_2}^{x_2} \right\}. \end{aligned} \quad (5.20)$$

Bằng cách triển khai bất đẳng thức Cauchy như trong [109, CT. (1.6)] cho hai giá trị thực dương $\bar{\varepsilon}_{\mathcal{U}_2}^{x_1}$ và $\bar{\varepsilon}_{\mathcal{U}_2}^{x_2}$, chúng ta có $\varepsilon_{\mathcal{U}_2}^{x_1} \varepsilon_{\mathcal{U}_2}^{x_2} \leq \frac{1}{2} (\varepsilon_{\mathcal{U}_2}^{x_1 2} + \varepsilon_{\mathcal{U}_2}^{x_2 2})$.

Tiếp theo, Chúng ta lấy trung bình $\mathbb{E} \left\{ \varepsilon_{\mathcal{U}_2}^{x_1} \varepsilon_{\mathcal{U}_2}^{x_2} \right\} \leq \frac{1}{2} \mathbb{E} \left[\varepsilon_{\mathcal{U}_2}^{x_1 2} + \varepsilon_{\mathcal{U}_2}^{x_2 2} \right]$. Chú ý rằng $\varepsilon_{\mathcal{U}_2}^{x_1}$ và $\varepsilon_{\mathcal{U}_2}^{x_2}$ có giá trị rất nhỏ (10^{-3} đến 10^{-5} [110]), do vậy nếu $\varepsilon_{\mathcal{U}_2}^{x_1} = \varepsilon_{\mathcal{U}_2}^{x_2}$ thì

$$\begin{aligned} \mathbb{E} \left\{ \varepsilon_{\mathcal{U}_2}^{x_1} \varepsilon_{\mathcal{U}_2}^{x_2} \right\} &\leq \mathbb{E} \left[\varepsilon_{\mathcal{U}_2}^{x_1 2} \right] \\ &< \mathbb{E} \left[\varepsilon_{\mathcal{U}_2}^{x_1} \right]. \end{aligned} \quad (5.21)$$

và

$$\begin{aligned}\mathbb{E}\left\{\varepsilon_{\mathcal{U}_2}^{x_1}\varepsilon_{\mathcal{U}_2}^{x_2}\right\} &\leq \mathbb{E}\left[\varepsilon_{\mathcal{U}_2}^{x_2^2}\right] \\ &< \mathbb{E}\left[\varepsilon_{\mathcal{U}_2}^{x_2}\right].\end{aligned}\tag{5.22}$$

Một cách tổng quát, chúng ta đã chứng minh được kết quả tương ứng với (5.19) dựa trên (5.21) và (5.22).

Bổ đề 2. Từ **Mệnh đề 1** và (5.15), BLER trung bình ở $\mathcal{U}_2$ cho quá trình giải mã tín hiệu x_1 và x_2 tương ứng, nhận được bên dưới như sau:

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_1}(x) = \begin{cases} v_2\sqrt{k}(\rho_{2H} - \rho_{2L}), & \frac{\alpha_1}{\alpha_2} < \rho_{2L}, \\ v_2\sqrt{k}\left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)}\mathcal{I}_3\right], & \rho_{2L} \leq \frac{\alpha_1}{\alpha_2} < \rho_{2H}, \\ v_2\sqrt{k}\left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)}\mathcal{I}_4\right], & \frac{\alpha_1}{\alpha_2} \geq \rho_{2H}, \end{cases}\tag{5.23}$$

và

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_2} = v_2\sqrt{k}\left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)}\left(\Delta(\sqrt{\rho_{2H}}) - \Delta(\sqrt{\rho_{2L}})\right)\right],\tag{5.24}$$

$$\begin{aligned}\Delta(x) &= \frac{1}{2}\left[x^2\Gamma\left(a+1, \frac{x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right) - \frac{x^{a+2}}{\left(b\sqrt{\alpha_2 G_2 \bar{\gamma}}\right)^a} e^{\frac{-x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}} \right. \\ &\quad \left. - \frac{(a+2)}{\left(\frac{1}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right)^2} \Gamma\left(a+2, \frac{x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right)\right].\end{aligned}\tag{5.25}$$

với

(Bổ đề 2 được chứng minh trong phụ lục F).

Trong trường hợp, $\bar{\gamma} \rightarrow \infty$, giá trị $\mathfrak{T}_L$ và $\mathfrak{T}_H$ trong (F.9) và (F.10) tiến đến bằng 0. Do đó, các tích phân trong (F.2) và (F.14) đều bằng 0. Như vậy, BLER trung bình tại $\mathcal{U}_2$ được viết gọn lại sau đây:

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_1} = v_2 \sqrt{k} (\rho_{2H} - \rho_{2L}). \quad (5.26)$$

5.3.3 Tỷ lệ lỗi khối dạng xấp xỉ

Chúng ta tính toán BLER tại $\mathcal{U}_i$ với sự tiếp cận gần đúng, tức là tìm biểu thức BLER dạng đóng cận trên khi SNR tiến đến vô cùng. Điều này có nghĩa là BLER của hệ thống luôn ở giá trị thấp hơn giá trị ngưỡng cận trên này, dù cho tỉ lệ SNR tăng cao đến mức nào đi nữa.

Để ý rằng giá trị $\rho_{iH} - \rho_{iL}$ trong (5.15) có thể thu được như sau:

$$\rho_{iH} - \rho_{iL} = \zeta_i + \frac{1}{2v_i \sqrt{k}} - \left(\zeta_i - \frac{1}{2v_i \sqrt{k}} \right) = \frac{1}{v_i \sqrt{k}}. \quad (5.27)$$

Bằng cách thay thế $v = [2\pi(2^{2\mathcal{R}_i} - 1)]^{-1/2}$ vào (5.27), $\rho_{iH} - \rho_{iL}$ rút gọn như sau:

$$\rho_{iH} - \rho_{iL} = \sqrt{2\pi(2^{2\mathcal{R}_i} - 1)} / k. \quad (5.28)$$

Bởi vì chiều dài khối tin ngắn với $k \geq 100$ CUs nên $\rho_{iH} - \rho_{iL}$ rất nhỏ. Hơn nữa, hàm $F_{\gamma_{\mathcal{U}_i}^x}(\gamma)$ là hàm xác định trong khoảng $[\rho_{iL}, \rho_{iH}]$. Bằng cách sử dụng xấp xỉ tích phân Riemann [71], BLER dạng xấp xỉ được dẫn ra như sau:

$$\tilde{\varepsilon}_{\mathcal{U}_i}^x \approx F_{\gamma_{\mathcal{U}_i}^x}(\zeta_i). \quad (5.29)$$

Hơn nữa, dựa trên xấp xỉ chuỗi vô hạn tương đương $1 - \exp\left(-\frac{x}{\bar{\gamma}}\right) \sim \frac{x}{\bar{\gamma}}$ cho (5.10)

và kẻ đó thay vào (5.29), BLER xấp xỉ ở $\mathcal{U}_i$ được dẫn ra như sau:

$$\begin{aligned}\tilde{\mathcal{E}}_{\mathcal{U}_1} &\simeq F_{\gamma_{\mathcal{U}_1}}(\zeta_1) \\ &\simeq \begin{cases} 1, & \zeta_1 \geq \frac{\alpha_1}{\alpha_2}, \\ \lambda_0 \frac{\zeta_1}{G_1 \bar{\gamma}(\alpha_1 - \alpha_2 \zeta_1)}, & \zeta_1 < \frac{\alpha_1}{\alpha_2}. \end{cases} \end{aligned} \quad (5.30)$$

Tiếp theo, bằng cách áp dụng [85, CT (8354.2)] cho (5.11) và (5.12), xấp xỉ hàm CDF của $\mathcal{U}_2$ được viết gọn lại sau đây:

$$F_{\gamma_{\mathcal{U}_2}^{\eta}}(x) \simeq \begin{cases} 1, & x \geq \frac{\alpha_1}{\alpha_2}, \\ \frac{1}{\Gamma(a+1)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{b} \sqrt{\frac{x}{G_2 \bar{\gamma}(\alpha_1 - \alpha_2 x)}} \right)^{a+1+m}}{m!(a+1+m)}, & x < \frac{\alpha_1}{\alpha_2}, \end{cases} \quad (5.31)$$

và

$$F_{\gamma_{\mathcal{U}_2}^{\chi}}(x) \simeq \frac{1}{\Gamma(a+1)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{b} \sqrt{\frac{x}{\alpha_2 G_2 \bar{\gamma}}} \right)^{a+1+m}}{m!(a+1+m)}, \quad (5.32)$$

với m là bậc của khai triển.

Ở chế độ SNR cao, xấp xỉ $F_{\gamma_{\mathcal{U}_2}^{\eta}}(x)$ và $F_{\gamma_{\mathcal{U}_2}^{\chi}}(x)$ có thể làm gọn bằng cách bỏ qua giá trị cao của m trong (5.31) và (5.32), tức là chọn $m = 0$. Từ đây, bằng cách kết hợp (5.27) với (5.29), xấp xỉ BLER ở $\mathcal{U}_2$ có thể thu được bên dưới như sau:

$$\tilde{\mathcal{E}}_{\mathcal{U}_2}^{\chi_1} \simeq F_{\gamma_{\mathcal{U}_2}^{\eta}}(\zeta_2) \simeq \begin{cases} 1, & \zeta_2 \geq \frac{\alpha_1}{\alpha_2}, \\ \frac{\left(\frac{1}{b} \sqrt{\frac{\zeta_2}{G_2 \bar{\gamma}(\alpha_1 - \alpha_2 \zeta_2)}} \right)^{a+1}}{\Gamma(a+1)(a+1)}, & \zeta_2 < \frac{\alpha_1}{\alpha_2}. \end{cases} \quad (5.33)$$

và

$$\tilde{\varepsilon}_{\mathcal{U}_2}^{x_2} \simeq F_{\gamma_{\mathcal{U}_2}^{x_2}}(\zeta_2) = \frac{\left(\left(\frac{1}{b} \sqrt{\frac{\zeta_2}{\alpha_2 G_2 \bar{\gamma}}}\right)\right)^{a+1}}{\Gamma(a+1)(a+1)}. \quad (5.34)$$

Cuối cùng, bằng cách thay thế (5.33) và (5.34) vào (5.19), chúng ta thu được xấp xỉ BLER tổng hợp tại $\mathcal{U}_2$ tương ứng như sau:

$$\tilde{\varepsilon}_{\mathcal{U}_2} = \tilde{\varepsilon}_{\mathcal{U}_2}^{x_1} + \tilde{\varepsilon}_{\mathcal{U}_2}^{x_2} - \min\{\tilde{\varepsilon}_{\mathcal{U}_2}^{x_1}, \tilde{\varepsilon}_{\mathcal{U}_2}^{x_2}\}. \quad (5.35)$$

5.3.4 Độ trễ và độ tin cậy

Độ trễ trung bình tương ứng khi truyền tín hiệu tới $\mathcal{U}_i$, đo bằng CUs, có thể được rút ra bên dưới như sau theo [51, CT (18)]:

$$\tau_i = \frac{k}{1 - \bar{\varepsilon}_{\mathcal{U}_i}}. \quad (5.36)$$

Tiếp theo, độ tin cậy trong quá trình truyền tín hiệu từ BS đến $\mathcal{U}_i$ được mô tả như xác suất mà một gói tin truyền từ nguồn BS đến $\mathcal{U}_i$ thành công. Độ tin cậy δ_i được suy ra như sau:

$$\delta_i = (1 - \bar{\varepsilon}_i) \cdot 100\%. \quad (5.37)$$

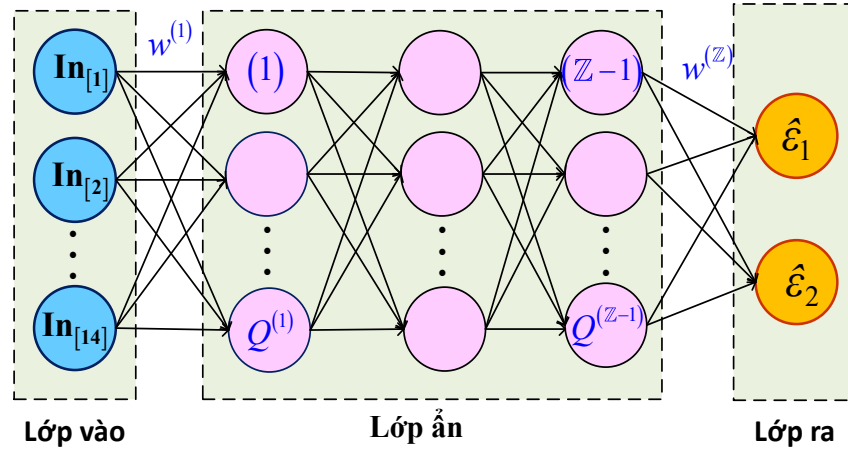
5.4 Mạng nơ ron sâu

Trong phần này, cấu trúc mạng DNN được thiết kế để dự đoán hiệu năng hệ thống thông qua BLER. Trong khi các phương pháp như Monte Carlo và phân tích số học đòi hỏi một lượng lớn thời gian chạy để vòng lặp hội tụ đến kết quả chính xác cao, điều này không phù hợp với các ứng dụng đòi hỏi thời gian thực (như mạng IoT). Mạng DNN dựa trên khả năng học sâu nên có khả năng dự đoán hiệu quả, không đòi hỏi tính toán phức tạp. Do đó, DNN được đề xuất để ước lượng đồng thời BLER của $\mathcal{U}_i$ một cách chính xác với thời gian xử lý nhanh và độ phức

tập tính toán thấp, nhằm thỏa mãn các yêu cầu của các kịch bản QoS khác nhau trong các mạng không dây hiện đại. Ngoài ra, chúng ta cũng xem xét ảnh hưởng của các yếu tố như chiều dài khối tin, các phân tử phản xạ lên BLER của hệ thống mạng.

5.4.1 Dự đoán BLER

Mô hình DNN được cấu thành từ một lớp đầu vào, $\mathbb{Z}-1$ các lớp ẩn, và một lớp đầu ra, được mô tả chi tiết trong Hình 5.2. Trong lớp đầu vào, có 14 nơ ron, trong đó mỗi nơ ron tương ứng với một giá trị tham số đầu vào được cung cấp trong Bảng 5.1. Trong $\mathbb{Z}-1$ lớp ẩn, mỗi lớp chứa $Q^{(z)}$ nơ ron với $z = 1, \dots, \mathbb{Z}-1$ và sử dụng hàm Sigmoid làm hàm kích hoạt [111].



Hình 5.2: Mô hình DNN cho dự đoán BLER tại $\mathcal{U}_i$, $i = \{1, 2\}$.

Hàm kích hoạt trong mạng DNN rất quan trọng vì giúp hỗ trợ trong việc học và nắm bắt các quan hệ phi tuyến và phức tạp giữa các đầu vào và ra tương ứng. Hàm Sigmoid, với tính phổ biến và dễ tính toán, được ưu tiên sử dụng trong mạng DNN được đề xuất.

Lớp ra có hai nơ ron tương ứng với BLER dự đoán của $\mathcal{U}_1$ và $\mathcal{U}_2$ (tương ứng với $\hat{\epsilon}_1$ và $\hat{\epsilon}_2$). Hơn nữa, chúng ta giả định lớp đầu ra kết nối hoàn toàn và không sử dụng các hàm kích hoạt.

Lớp đầu ra, $y^{(z)}$ được tính toán như [112, CT. (13)]:

$$y^{(z)} = g^{(z)} \left(w^{(z)} y^{(z-1)} + b^{(z)} \right), \quad (5.38)$$

với $g^{(z)}(x) = \frac{1}{1+e^{-x}}$ là hàm kích hoạt Sigmoid, $w^{(z)}(.)$ và $b^{(z)}(.)$ là lớp của trọng số và độ lệch.

Bảng 5.1: Tham số tính BLER.

Số thứ tự	Tham số	Giá trị	Số thứ tự	Tham số	Giá trị
1	G_1	1	8	k	512(CUs)
2	G_2	0.1	9	η	2
3	$\bar{\gamma}$	[0dB,20dB]	10	β_1	256bits
4	N	2	11	β_2	256bits
5	α_1	[0.5;0.6;0.8]	12	d_1	1
6	α_2	[0.5;0.4;0.2]	13	d_2	1
7	L	4	14	d_0	1

Dựa vào (5.38), giá trị dự đoán của $\hat{\varepsilon}_i$ có thể được trình bày như sau:

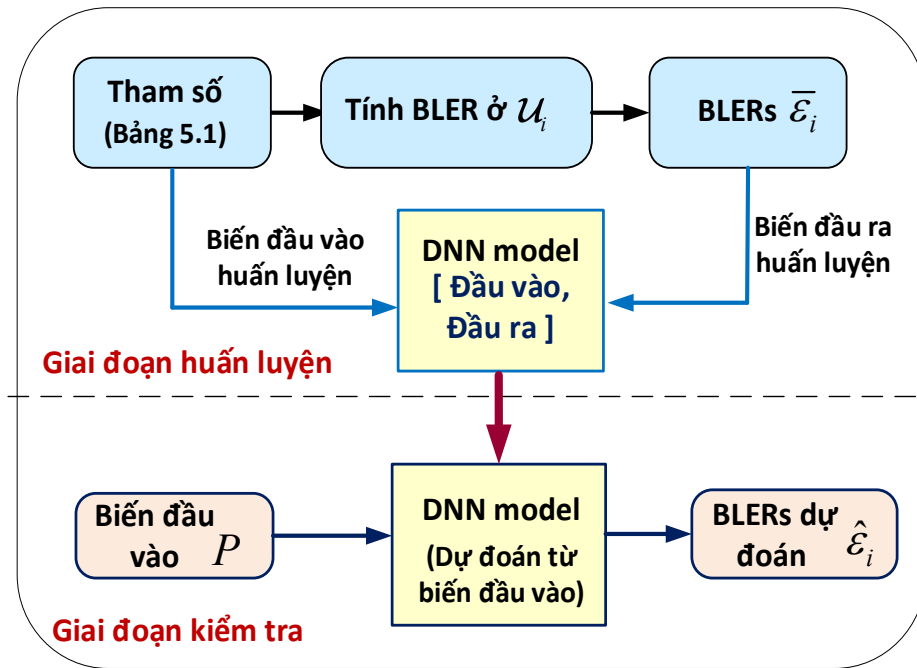
$$\hat{\varepsilon}_i = \mathcal{G}(\mathcal{IN}; \Omega), \quad (5.39)$$

với $\mathcal{G}(.)$ là hàm ánh xạ của các giá trị đầu vào $\mathcal{IN}$ và Ω , tương ứng và có thể tìm được từ quá trình học thông qua mạng DNN. Do đó, chúng ta đặt $\Omega \triangleq \{w^{(z)}, b^{(z)}, z = 1, \dots, \mathbb{Z}-1\}$ là tập hợp trọng số và độ lệch. Tiếp theo, mạng DNN được chia thành hai giai đoạn là huấn luyện và dự đoán như trong Hình 5.3.

Trong giai đoạn huấn luyện, thuật toán Levenberg–Marquardt được áp dụng để cải thiện các tham số của mô hình, dựa trên bộ dữ liệu qua quá trình học ngoại

tuyến. LMA nổi tiếng với hiệu suất cao, là lựa chọn ưa thích cho việc huấn luyện mạng DNN. Tùy thuộc vào tốc độ học trong quá trình lan truyền ngược, trọng số và độ lệch được cập nhật lặp lại. Vì toàn bộ giai đoạn huấn luyện được thực hiện ngoại tuyến, mạng được thiết kế có thể giảm thiểu độ phức tạp tính toán, chi phí triển khai và thời gian thực hiện [106].

Sau khi quá trình huấn luyện được hoàn thành, mạng DNN được tạo ra, trong đó các kết quả đạt được có thể được sử dụng cho giai đoạn dự đoán trực tuyến bất cứ khi nào có dữ liệu đầu vào mới.



Hình 5.3: Giai đoạn huấn luyện và kiểm tra của mạng DNN.

d. Tạo tập dữ liệu

Để xây dựng một bộ dữ liệu để huấn luyện cho mạng DNN, tính toán BLER trung bình tại U_i , thông qua sử dụng các tham số được mô tả như trong Bảng 5.1 và áp dụng các công thức trong (5.16) và (5.19). Những tham số này được trích xuất làm các biến đầu vào, còn $\bar{\varepsilon}_i$ là biến đầu ra cho các mẫu tương ứng. Nói chung, trong tập dữ liệu được tạo $\mathcal{M}$, mỗi mẫu j là một vec tơ hàng tương ứng $[In_{[j]}, Out_{[j]}]$, $j = 1, \dots, \mathcal{M}$ với $In_{[j]}$ và $Out_{[j]}$ là các giá trị của biến đầu vào và đầu

ra, tương ứng. Từ cơ sở này, tập $\mathcal{M}$ được tạo thành với tổng cộng 2000 mẫu, chia ngẫu nhiên thành tập huấn luyện và tập kiểm tra để mô hình dự đoán học và kiểm tra như được mô tả trong Hình 5.3. Cụ thể, 70% được phân bổ cho tập huấn luyện $\mathcal{M}_{\text{train}}$, trong khi phần còn lại 30% của $\mathcal{M}$ được chia đều cho tập đánh giá $\mathcal{M}_{\text{val}}$ và tập kiểm tra $\mathcal{M}_{\text{tes}}$. Cụ thể, chúng ta giả định rằng hệ thống mạng đề xuất được đặt trong không gian phẳng hai chiều. Do đó, các khoảng cách chuẩn hóa của nguồn BS, IRS, $\mathcal{U}_1$ và $\mathcal{U}_2$ được xác định như một hàm của vị trí như $d_{AB} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$ với (x_A, x_B) chỉ vị trí của nút nguồn BS, IRS, $\mathcal{U}_1$ và $\mathcal{U}_2$, tương ứng. Ở đây, tọa độ của các nút được đặt thành $(0,0)$, $(1,0)$, $(0,1)$ và $(2,0)$ tương ứng với những khoảng cách của d_0 , d_1 và d_2 được trình bày như trong Bảng 5.1.

e. Chỉ số đánh giá

Trong phần này, chúng ta đánh giá tính hiệu quả của mô hình mạng DNN đề xuất dựa vào hai chỉ số thống kê như sau: căn bậc hai sai số bình phương trung bình (RMSE) và sai số tuyệt đối trung bình phần trăm (MAPE). Các chỉ số thống kê này được tính như bên dưới:

$$\text{RMSE} = \sqrt{\frac{1}{\mathcal{M}_{\text{tes}}} \sum_{j=1}^{\mathcal{M}_{\text{tes}}} \left(\bar{\varepsilon}_i^{(j)} - \hat{\varepsilon}_i^{(j)} \right)^2}, \quad (5.40)$$

và

$$\text{MAPE} = \frac{1}{\mathcal{M}_{\text{tes}}} \sum_{j=1}^{\mathcal{M}_{\text{tes}}} \frac{\left| \bar{\varepsilon}_i^{(j)} - \hat{\varepsilon}_i^{(j)} \right|}{\bar{\varepsilon}_i^{(j)}}, \quad (5.41)$$

với $\bar{\varepsilon}_i^{(j)}$ và $\hat{\varepsilon}_i^{(j)}$ lần lượt là BLER huấn luyện ban đầu và BLER được dự đoán từ mô hình DNN, tương ứng.

RMSE và MAPE đều là những phương pháp hiệu quả để đo lường sự chênh lệch giữa các giá trị dự đoán và thực tế. Hơn nữa, chúng ta lưu ý rằng cả hai phương

pháp này đều tỷ lệ nghịch với số lượng mẫu M_{tes} , điều đó có nghĩa là sai số càng nhỏ, số lượng M_{tes} càng lớn. Nói chung, giá trị sai số tốt nhất cho mô hình hồi quy là số không. Cho đến nay, vẫn chưa có sự đồng thuận về một chỉ số chung để đo lường hiệu quả của các mô hình hồi quy. Do đó, RMSE và MAPE được chọn ngẫu nhiên trong số các phương pháp hiệu quả khác cho mô hình đề xuất. Hơn nữa, quan sát (5.40) và (5.41) chúng ta thấy rõ MAPE cho kết quả nhỏ hơn so với RMSE [113].

f. Độ phức tạp tính toán

Độ phức tạp tính toán của mô hình mạng DNN thiết kế được đo bằng tổng số tham số tính toán ($Para$) và số phép tính dấu phẩy động trên giây (FLOPs), tức tổng số phép toán cần thực hiện trong mạng [114]. Cụ thể, tổng số tham số và số FLOPs được tính toán lần lượt như sau:

$$Para = \underbrace{L_{in} + \sum_{\ell=1}^{L_{hid}} L_{neu}^{(\ell)} + L_{out}}_{\text{tong so nơ ron}} + \underbrace{\sum_{\ell=1}^{L_{hid}} L_{neu}^{(\ell)}}_{\text{tong do lech}} + \underbrace{L_{in} L_{neu}^{(1)} + \sum_{\ell=1}^{L_{hid}-1} L_{neu}^{(\ell)} L_{neu}^{(\ell+1)} + L_{neu}^{(L_{hid})} L_{out}}_{\text{tong trong so}}, \quad (5.42)$$

và

$$FLOP = 2 \cdot \left\{ L_{in} L_{neu}^{(1)} + \sum_{\ell=1}^{L_{hid}-1} L_{neu}^{(\ell)} L_{neu}^{(\ell+1)} + L_{neu}^{(L_{hid})} L_{out} \right\}, \quad (5.43)$$

với L_{in} , L_{out} , và L_{in} là số lượng nơ ron ở lớp đầu vào, lớp đầu ra và ở lớp ẩn thứ ℓ , tương ứng.

Tổng số tham số $Para$ và FLOP thường được tối ưu tùy theo yêu cầu của từng ứng dụng cụ thể. Đối với các hệ thống yêu cầu tính toán nhanh và triển khai trên các thiết bị có tài nguyên hạn chế, như thiết bị IoT (dưới 1 triệu tham số) hoặc thiết bị di động (dưới 1 GFLOP), việc giảm thiểu tham số và FLOP là cần thiết. Ngược lại, trong các ứng dụng đòi hỏi độ chính xác cao và có sự hỗ trợ từ phần cứng mạnh mẽ, mức độ phức tạp của mô hình có thể được gia tăng đáng kể để đạt hiệu quả tối ưu.

5.4.2 Dự đoán công suất

Mục tiêu chính của hệ thống đề xuất là đáp ứng các yêu cầu QoS của uRLLCs, nghĩa là BLER của $\mathcal{U}_i$ phải nhỏ hơn giá trị giới hạn 10^{-5} [7] và đòi hỏi SNR ở mức cao. Vì BLER tại $\mathcal{U}_i$ của hệ thống khá phức tạp nên việc tính toán trực tiếp đòi hỏi nhiều thời gian. Do đó, vấn đề tối ưu phân chia công suất được xem xét thông qua sử dụng xấp xỉ của BLER và được mô tả như sau:

$$\begin{aligned} \min_{\mathbf{P}_i} \quad & \bar{\varepsilon}_i(\mathbf{P}_i) \\ \text{s.t.} \quad & \alpha_1 \mathbf{P}_1 + \alpha_2 \mathbf{P}_2 \leq P_0, \\ & \alpha_1 + \alpha_2 = 1, \end{aligned} \quad (5.44)$$

với P_0 là tổng nguồn công suất phân bổ của hệ thống.

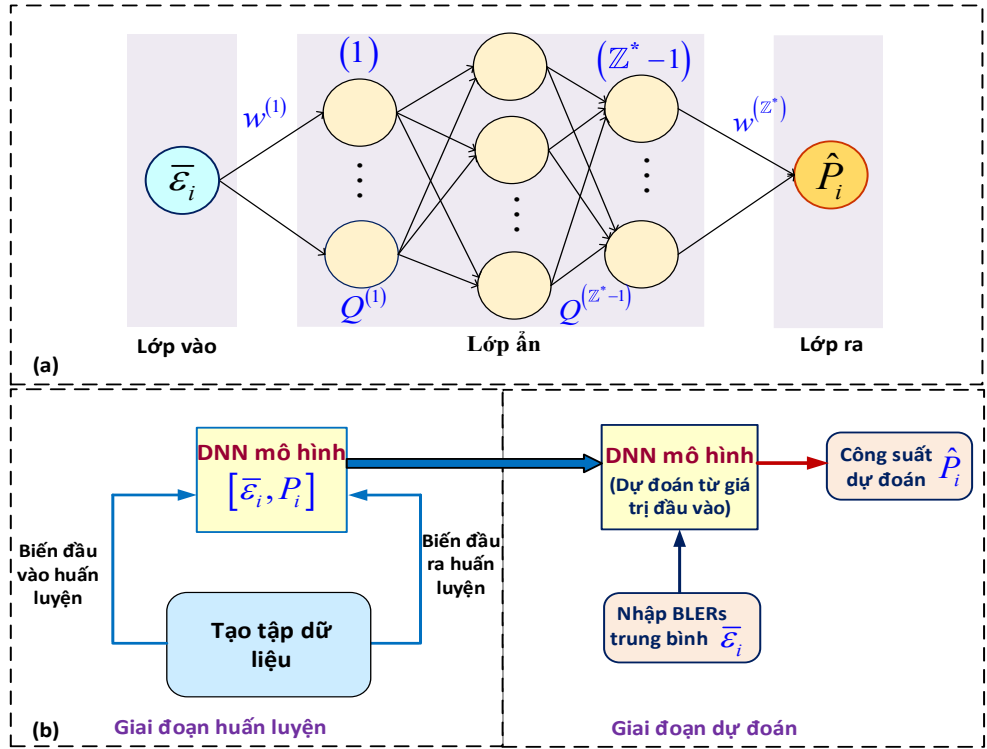
Từ (5.16) và (5.19), hàm mục tiêu $\bar{\varepsilon}_i(P_i)$ là hàm không lồi nên vấn đề tối ưu trong (5.44) là vấn đề không lồi. Điều này dẫn đến vấn đề tối ưu trong (5.44) khó có thể thực hiện bằng phương pháp đạo hàm riêng. Để giải quyết thử thách này, mô hình DNN được thiết kế để dự đoán công suất phát P_i của $\mathcal{U}_i$ theo các yêu cầu của uRLLCs.

Tương tự như việc dự đoán BLER, trong phần này thiết kế của mô hình DNN được tiến hành thông qua quá trình huấn luyện và kiểm thử như được minh họa trong Hình 5.4. Mô hình cũng bao gồm một lớp đầu vào (chứa một nơ ron $\bar{\varepsilon}_i$), $\mathbb{Z}^* - 1$ lớp ẩn với $Q^{(z^*)}$ nơ ron trên mỗi lớp với $z^* = 1, \dots, \mathbb{Z}^* - 1$ và một lớp đầu ra (chứa một nơ ron đầu ra $\hat{P}_i$), tương ứng. Trong mô hình này, hàm Sigmoid được sử dụng và thuật toán LMA như trong phần dự đoán BLER.

Giá trị đầu ra dự đoán cho DNN với hàm ánh xạ $\mathcal{G}^*(.)$, các giá trị đầu vào $\bar{\varepsilon}_i$ và tập trọng số và độ lệch $\Omega^* \triangleq \left\{ w^{(z^*)}, b^{(z^*)}, z^* = 1, \dots, \mathbb{Z}^* - 1 \right\}$, tương ứng, được tính toán như sau:

$$\hat{P}_i = \mathcal{G}^*(\bar{\varepsilon}_i; \Omega^*). \quad (5.45)$$

Ngoài ra, mô hình dự đoán sử dụng các tham số như được mô tả trong Bảng 5.1. Trong đó, một trong các hệ số phân chia công suất, như là $\alpha_1 = 0.6$ và $\alpha_2 = 0.4$ được sử dụng để tính BLER tại $\mathcal{U}_i$, dựa vào (5.16) và (5.19). Tiếp theo, $\bar{\varepsilon}_i \leq 10^{-5}$ được chọn làm biến đầu vào và đầu ra P_i để thực hiện hai giai đoạn huấn luyện và dự đoán như được minh họa trong Hình 5.4. Bộ dữ liệu được tạo ra $\mathcal{M}^*$ được mô tả bằng các giá trị của vector đặc trưng cả đầu vào và đầu ra, tức là $[\varepsilon^{(j*)}, P^{(j*)}]$, $j^* = 1, \dots, \mathcal{M}^*$. Kết quả là, có 2000 mẫu được tạo ra và cũng được chia thành tỷ lệ 80% cho tập huấn luyện, $\mathcal{M}_{\text{train}}^*$, 20% của $\mathcal{M}^*$ chia đều cho tập đánh giá $\mathcal{M}_{\text{val}}^*$ và tập kiểm tra $\mathcal{M}_{\text{tes}}^*$.



Hình 5.4: Mô hình DNN cho dự đoán công suất tại $\mathcal{U}_i$.

5.5 Mô phỏng và thảo luận

Trong phần này, hiệu năng của hệ thống đề xuất được xác minh thông qua mô phỏng Monte Carlo. Hơn nữa, hiệu quả mạng DNN trong dự đoán BLER và dự đoán công suất phát của hệ thống được chứng minh bằng cách thực hiện trên phần mềm Matlab.

5.5.1 Đánh giá hiệu năng

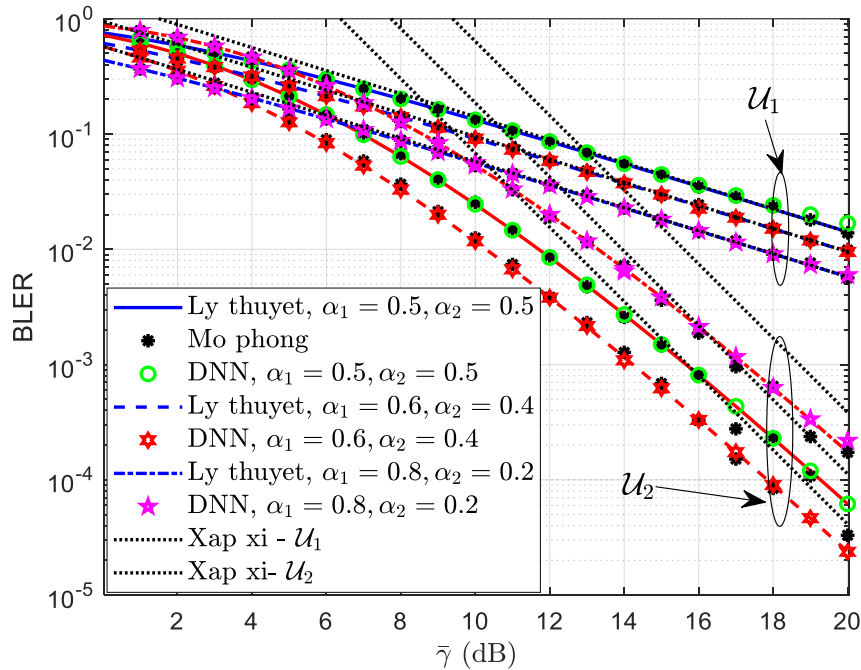
Để thuận lợi cho quá trình mô phỏng, chúng ta sử dụng các tham số mô phỏng chính trong toàn bộ mục này, được mô tả trong Bảng 5.1 chi tiết dưới đây: $G_1 = 1$, $G_2 = 0.1$, $\bar{\gamma} = 20$ dB, $k = 512$ CUs, $\eta = 2$ [86], $\beta_1 = \beta_2 = 256$ bits [115], $N = 2$, $\alpha_1 = [0.5; 0.6; 0.8]$, $L = 4$, $\alpha_2 = [0.5; 0.4; 0.2]$, và $d_0 = d_1 = d_2 = 1$. Về mô phỏng DNN, các tham số huấn luyện được minh họa như trong Bảng 5.2. Đầu tiên, Hình 5.5 trình bày ảnh hưởng của BLER vào SNR trung bình $\bar{\gamma}$ tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với ba kịch bản của hệ số phân chia công suất như $\alpha_1 = \alpha_2 = 0.5$, $\alpha_1 = 0.6, \alpha_2 = 0.4$ và $\alpha_1 = 0.8, \alpha_2 = 0.2$, tương ứng. Quan sát Hình 5.5, chúng ta thấy rằng BLER của $\mathcal{U}_1$ và $\mathcal{U}_2$ giảm trong khoảng $\bar{\gamma}$ ở mức trung bình và cao theo các hệ số phân bổ công suất α_1 và α_2 của $\mathcal{U}_1$ và $\mathcal{U}_2$. Đặc biệt, BLER ở $\mathcal{U}_2$ luôn thấp hơn BLER ở $\mathcal{U}_1$, tức là hiệu năng ở $\mathcal{U}_2$ cải thiện hơn mặc dù phân bổ công suất thấp hơn.

Bảng 5.2: Tham số huấn luyện DNN cho dự đoán BLER.

Tham số huấn luyện	Giá trị
Số lớp ẩn	4 lớp
Số nơ ron trong các lớp ẩn	100 nơ ron
Tốc độ học	10^{-3}
Hàm kích hoạt	Sigmoid
Epoch cho giai đoạn huấn luyện	1000 <i>epochs</i>

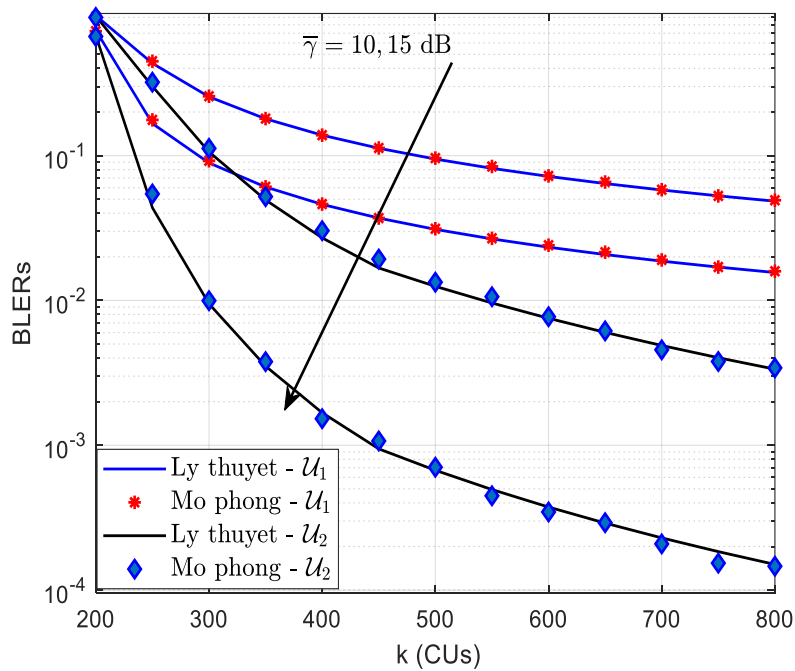
Hơn nữa, chúng ta cũng thấy rõ sự ảnh hưởng của BLER của $\mathcal{U}_1$ và $\mathcal{U}_2$ khi các giá trị α_1 và α_2 tăng hoặc giảm. Cụ thể, BLER của $\mathcal{U}_1$ càng giảm khi α_1 càng tăng. Ngược lại đối với $\mathcal{U}_2$, BLER ở $\mathcal{U}_2$ thấp nhất đạt được khi giá trị của $\alpha_2 = 0.4$.

Tuy nhiên, khi $\alpha_1 = \alpha_2 = 0.5$ BLER của $\mathcal{U}_2$ giảm nhưng BLER của $\mathcal{U}_1$ đạt giá trị cao nhất. Do đó, để đảm bảo cung cấp năng lượng cho cả $\mathcal{U}_1$ và $\mathcal{U}_2$, chúng ta chọn hệ số $\alpha_1 = 0.6$ và $\alpha_2 = 0.4$ cho tất cả mô phỏng ở các hình trong phần sau. Ngoài ra, chúng ta quan sát thấy kết quả dự đoán bằng DNN (đường hình sao) trùng khớp với các đường kết quả phân tích và kết quả mô phỏng ở mọi giá trị $\bar{\gamma}$ tại $\mathcal{U}_1$ và $\mathcal{U}_2$. Ở vùng $\bar{\gamma}$ cao, đường mô phỏng không trùng tuyệt đối với hai đường kết quả lý thuyết và mô phỏng DNN (do đường mô phỏng BLER của $\mathcal{U}_2$ lấy kết quả từ công thức dạng xấp xỉ (5.35)). Điều này là bước đầu cho chúng ta thấy lợi ích của DNN, và sự cải tiến của DNN sẽ được phân tích chi tiết hơn qua Hình 5.9 và Bảng 5.3.



Hình 5.5: Sự tác động của SNR trung bình lên BLER tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với ba trường hợp hệ số $\alpha_1 = \alpha_2 = 0.5$, $\alpha_1 = 0.6, \alpha_2 = 0.4$ và $\alpha_1 = 0.8, \alpha_2 = 0.2$.

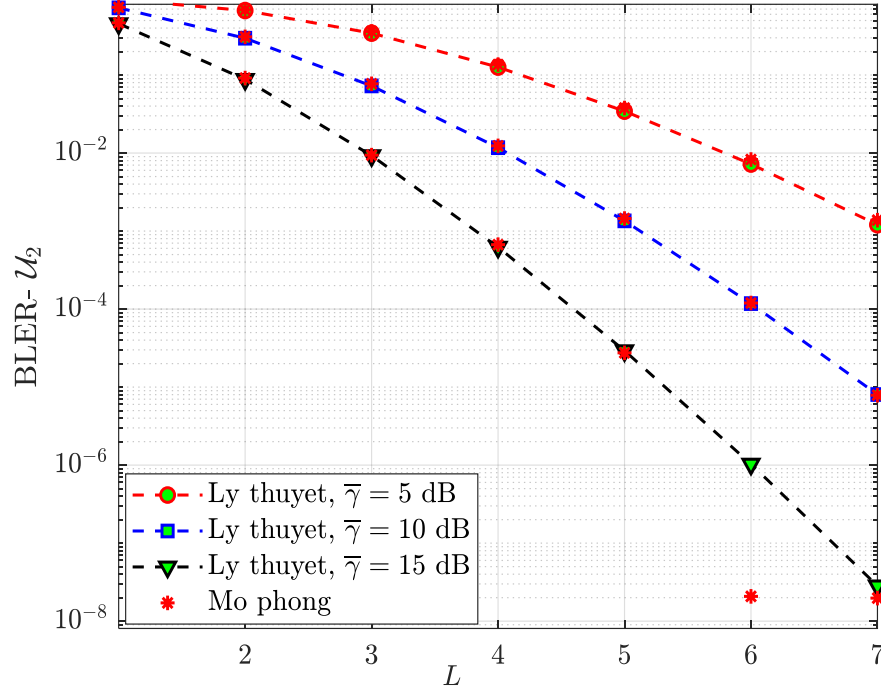
Tiếp theo, Hình 5.6 trình bày ảnh hưởng của k vào BLER trung bình tại $\mathcal{U}_1$ và $\mathcal{U}_2$, tương ứng. Như quan sát Hình 5.6, BLER trung bình tại $\mathcal{U}_1$ và $\mathcal{U}_2$ giảm khi k tăng lên ở hai trường hợp $\bar{\gamma}$ là $\bar{\gamma} = 15$ dB và $\bar{\gamma} = 20$ dB. Chúng ta để ý rằng, BLER trung bình của $\mathcal{U}_2$ nhỏ hơn BLER trung bình của $\mathcal{U}_1$ ở tất cả giá trị $\bar{\gamma}$. Điều này làm nổi bật sự cải thiện hiệu năng của $\mathcal{U}_2$ so với hiệu năng của $\mathcal{U}_1$, mặc dù hệ số phân bổ công suất của $\mathcal{U}_2$ thấp hơn. Tuy nhiên, chúng ta cũng cần xem xét đến sự đánh đổi giữa k lớn và độ trễ lớn qua Hình 5.8. Ví dụ, để đảm bảo yêu cầu dịch vụ uRLLCs BLER của $\mathcal{U}_2$ phải đạt $\leq 10^{-5}$ với $\bar{\gamma} = 20$ dB, chiều dài khối k được chọn là 600 CUs.



Hình 5.6: Sự tác động của k lên BLER tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với $\alpha_1 = 0.6$ và $\alpha_2 = 0.4$.

Tiếp theo, Hình 5.7 mô tả ảnh hưởng của các thành phần phản xạ L lên BLER trung bình của $\mathcal{U}_2$ với ba trường hợp của $\bar{\gamma}$ như 5 dB, 10 dB và 15 dB, tương ứng. Như được quan sát hình vẽ, BLER trung bình của $\mathcal{U}_2$ giảm đáng kể khi tăng giá trị của L và $\bar{\gamma}$. Bởi vì việc tăng $\bar{\gamma}$ có giới hạn nên chúng ta có thể tăng L phù

hợp để cải thiện hiệu năng của $\mathcal{U}_2$. Tiếp theo, Hình 5.8 mô tả sự ảnh hưởng của độ trễ và độ tin cậy vào chiều dài khối k tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với hai trường hợp của $\bar{\gamma}$ cụ thể $\bar{\gamma} = 15$ dB và $\bar{\gamma} = 20$ dB, tương ứng.

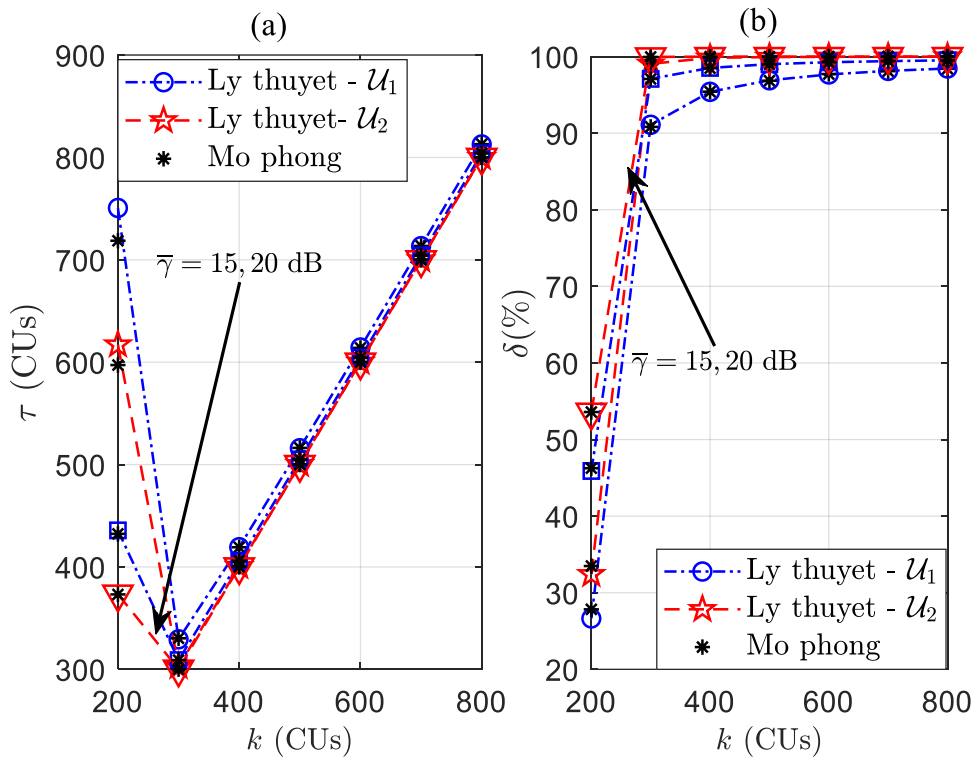


Hình 5.7: Sự tác động của L lên BLER tại $\mathcal{U}_2$ với ba trường hợp $\bar{\gamma} = 5$ dB, $\bar{\gamma} = 10$ dB và $\bar{\gamma} = 15$ dB.

Hình 5.8 thấy rằng độ trễ và độ tin cậy của $\mathcal{U}_2$ cải thiện hơn so với $\mathcal{U}_1$ tất cả trường hợp của $\bar{\gamma}$. Chúng ta để ý rằng khi k tăng dẫn đến độ trễ cũng tăng lên tương ứng ở cả $\mathcal{U}_1$ và $\mathcal{U}_2$. Đặc biệt, ở $k = 300$ CUs, độ trễ của cả hai người dùng đạt giá trị tốt nhất và độ tin cậy gần như bão hòa khi $k = 400$ CUs, ví dụ ở $\mathcal{U}_2$ độ trễ $\tau = 300$ CUs với $\bar{\gamma} = 20$ dB. Điều này hợp lý với các kết quả mô phỏng được minh họa như trong Hình 5.6 do sự giảm chậm của BLER từ $k = 400$ CUs. Do đó, giả sử nếu chúng ta chọn $k = 600$ CUs (xét trong khoảng thời gian CUs $\approx 3 \mu s$),

hệ thống sẽ phải chịu độ trễ là $\tau = 1.8 \text{ ms}$. Độ trễ này đáp ứng yêu cầu về độ trễ của uRLLCs để có thể phục vụ cho mạng Internet cảm ứng [116].

Tiếp theo, Quan sát Hình 5.9, chúng ta có thể thấy được rằng mô hình DNN mang lại kết quả dự đoán với độ tin cậy cao (RMSE và MAPE nhỏ hơn 10^{-5}). Khi tăng số lượng mẫu kiểm tra thì kết quả càng đáng tin cậy. Khi số lượng mẫu kiểm tra tăng, độ chính xác của dự đoán cũng được cải thiện đáng kể. Kết quả này hoàn toàn phù hợp với các công thức lý thuyết (5.40 và 5.41).



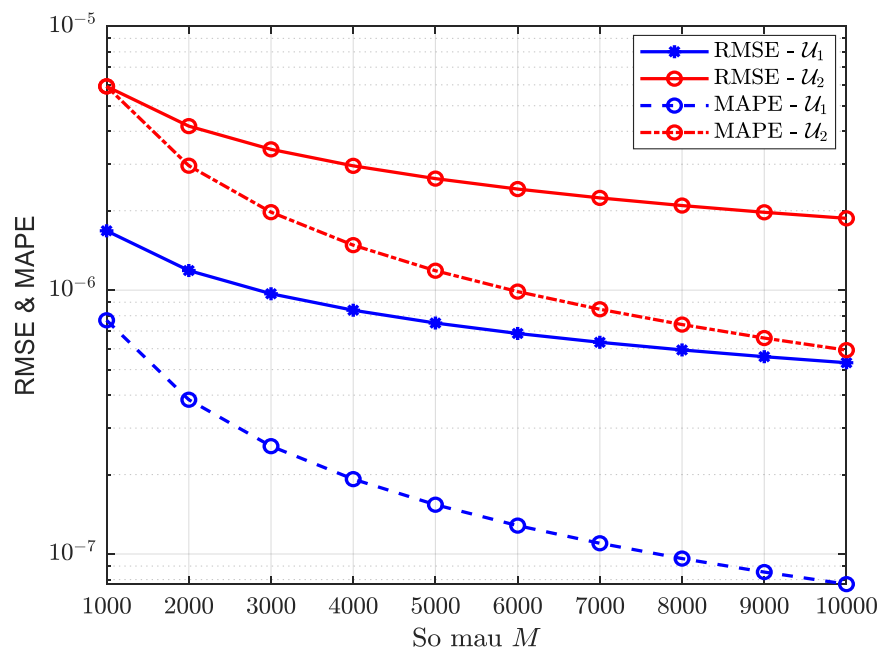
Hình 5.8: Ảnh hưởng độ trễ và độ tin cậy lên k tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với trường hợp $\bar{\gamma} = 15 \text{ dB}$ và $\bar{\gamma} = 20 \text{ dB}$.

Hơn nữa, như minh họa trong Bảng 5.3, phương pháp DNN vượt trội về thời gian thực thi so với các phương pháp truyền thống, đặc biệt trong các bài toán tính toán phức tạp. Tốc độ xử lý nhanh của giai đoạn trực tuyến trong DNN là một lợi thế rõ, vượt xa so với các giai đoạn ngoại tuyến hoặc các phương pháp truyền thống như mô phỏng và phân tích lý thuyết. Điều này khẳng định tiềm năng ứng dụng

mạnh mẽ của DNN trong các bài toán thực tế, và yêu cầu cả độ chính xác cao và thời gian xử lý nhanh.

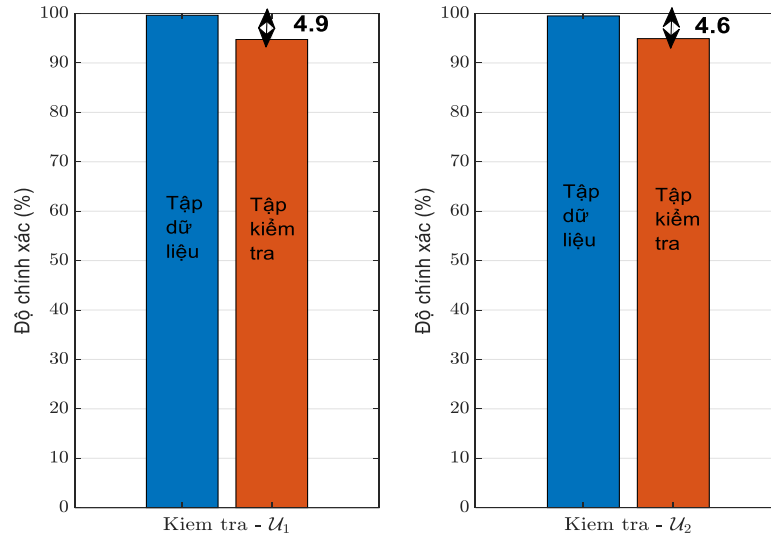
Bảng 5.3: So sánh thời gian thực thi hệ thống trong ba phương pháp: phân tích lý thuyết, mô phỏng và mạng DNN tại $\mathcal{U}_1$ và $\mathcal{U}_2$.

Trường hợp	Thời gian thực hiện (giây)
Lý thuyết - $\mathcal{U}_1$	1.336408
Lý thuyết - $\mathcal{U}_2$	6.390332
Lý thuyết - $\mathcal{U}_1$	5.629326
Mô phỏng - $\mathcal{U}_2$	25.463482
DNN- $\mathcal{U}_1$	0.002797
DNN- $\mathcal{U}_2$	0.002962



Hình 5.9: RMSE và MAPE so với số lượng mẫu kiểm tra.

Cuối cùng, Hình 5.10 được sử dụng để đánh giá mức độ tin cậy của mô hình DNN khi áp dụng trên hai kịch bản kiểm tra tại $\mathcal{U}_1$ và $\mathcal{U}_2$ với dữ liệu kiểm tra bao gồm cả những mẫu từng xuất hiện trong quá trình huấn luyện (thuộc tập huấn luyện) và những mẫu được chọn ngẫu nhiên.



Hình 5.10: So sánh phần trăm độ chính xác của kết quả kiểm tra mô hình DNN tại $\mathcal{U}_1$ và $\mathcal{U}_2$.

Độ chính xác của mô hình được tính theo công thức như $100 \times \frac{1}{n} \sum_{i=1}^n \frac{|y_i - y_i^{\text{tes}}|}{y_i}$

, trong đó y_i là giá trị dự đoán trên tập huấn luyện và y_i^{tes} là giá trị dự đoán trên tập kiểm tra, n là số mẫu.

Quan sát Hình 5.10, chúng ta nhận thấy rằng độ chính xác trên tập kiểm tra thuộc tập dữ liệu huấn luyện ban đầu đối với $\mathcal{U}_1$ đạt 99,6%, trong khi độ chính xác trên tập kiểm tra được chọn ngẫu nhiên là 94,7%. Mức chênh lệch giữa hai giá trị này vào khoảng 4,9%. Tương tự, trong trường hợp $\mathcal{U}_2$, độ chênh lệch là khoảng 4,6%. Hai giá trị sai lệch này, đều dưới ngưỡng 5%, cho thấy mô hình DNN có khả năng tổng quát hóa tương đối tốt và có thể chấp nhận được trong bối cảnh dữ liệu ít biến động và hệ thống không quá phức tạp.

Tuy nhiên, nếu mục tiêu của mô hình là triển khai trong môi trường mạng thực tế, nơi điều kiện hoạt động có thể thay đổi lớn và không thể dự đoán trước, thì độ chính xác cần phải ở mức cao hơn, và mức sai lệch gần 5% này là không đảm bảo. Trong trường hợp này, chúng ta phải huấn luyện lại mô hình DNN với dữ liệu theo điều kiện thực tế nhằm nâng cao tính ổn định và độ tin cậy của mô hình dự đoán.

5.5.2 Dự đoán công suất

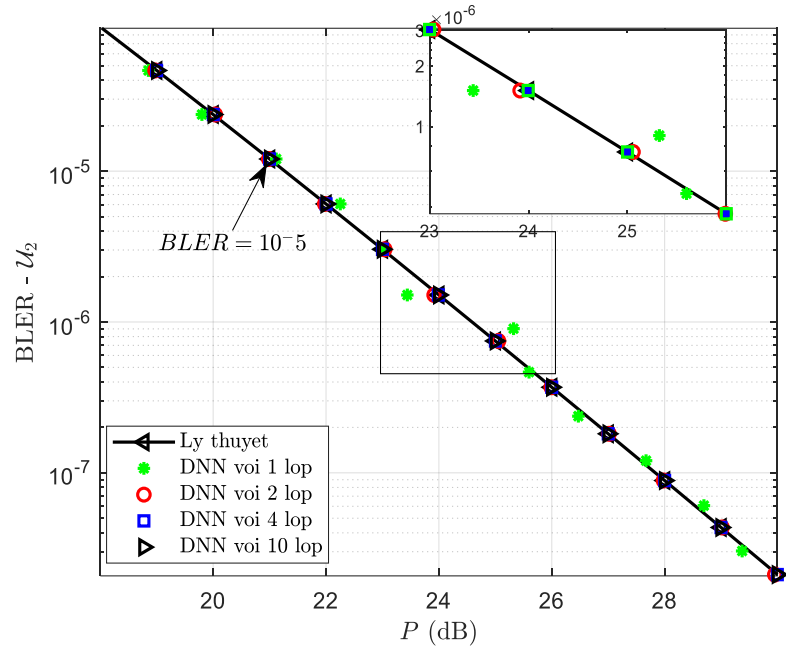
Trong phần này, kết quả dự đoán công suất phân bố cho $\mathcal{U}_2$ được trình bày dựa trên mô hình DNN. Các bước tiến hành của DNN được thực hiện dựa trên phần mềm Matlab và các thông số huấn luyện được cho tương tự như phần A (*Dự đoán BLER*), tức là số nơ ron trên mỗi lớp là 100 nơ-ron, epoch cho giai đoạn huấn luyện là 1000 epochs, tốc độ học là 10^{-3} và hàm kích hoạt Sigmoid cũng được áp dụng như mô tả trong Bảng 5.2. Đặc biệt, các giá trị đầu vào của BLER trung bình cho giai đoạn huấn luyện được chọn $\leq 10^{-5}$. Hơn nữa, trong trường hợp này, tập dữ liệu được tạo ra với khoảng 2000 mẫu, được phân chia thành ba phần với tỉ lệ 80% dành cho huấn luyện, 20% cho việc kiểm tra, và 20% cho đánh giá.

Tiếp theo, chúng ta trình bày ảnh hưởng của BLER trung bình của $\mathcal{U}_2$ vào công suất phát P , với các lớp ẩn khác nhau như trong Hình 5.11.

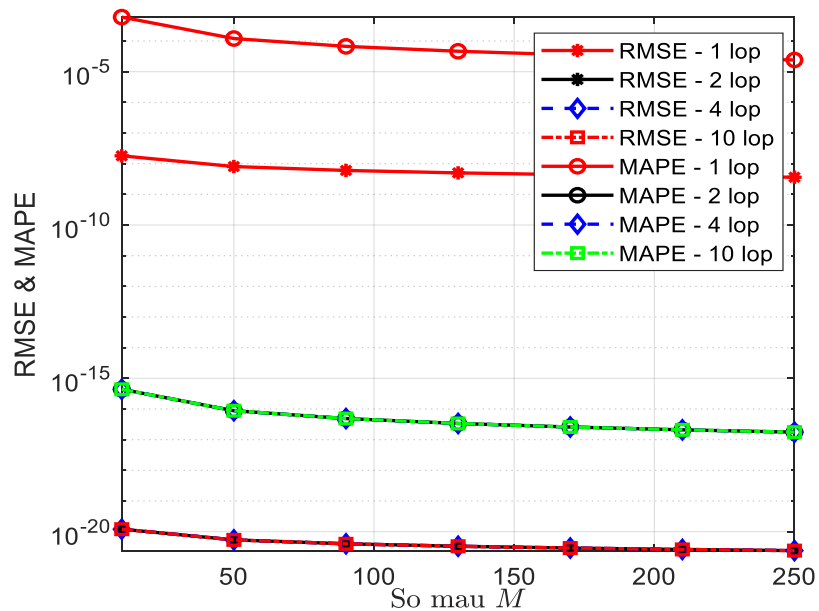
Như đã quan sát, BLER trung bình của $\mathcal{U}_2$ giảm đáng kể khi P tăng đối với tất cả các lớp ẩn. Cụ thể, các kết quả dự đoán bằng mô hình DNN hoàn toàn khớp với kết quả mong muốn (đường liền) ở 2 lớp ẩn trở lên. Điều này dẫn đến sai số RMSE và sai số MAPE thấp hơn trong kết quả dự đoán của công suất P như trong Hình 5.12.

Chúng ta quan sát thấy rằng, số lớp ẩn trong mô hình DNN có ảnh hưởng đáng kể đến độ chính xác của dự đoán. Cụ thể, khi số lớp ẩn tăng, giá trị RMSE và MAPE của dự đoán giảm, đồng nghĩa với việc độ tin cậy tăng lên. Điều này có thể giải thích bởi việc các mô hình DNN với nhiều lớp ẩn có số lượng trọng số lớn

hơn, từ đó cải thiện khả năng học và khái quát hóa dữ liệu. Trái lại, các mô hình chỉ có một lớp ẩn thường bị giới hạn trong việc học các đặc trưng phức tạp từ dữ liệu đa chiều, dẫn đến độ chính xác thấp hơn.



Hình 5.11: BLER so với công suất phân bổ tại $\mathcal{U}_2$ ở các lớp khác nhau.



Hình 5.12: Ảnh hưởng RMSE và MAPE của DNN vào số lớp ẩn và số mẫu.

Tuy nhiên, việc tăng số lớp ẩn không chỉ cải thiện hiệu suất, mà còn kéo theo sự gia tăng đáng kể về độ phức tạp tính toán. Theo công thức (5.42) và (5.43) . nếu mô hình có 4 lớp ẩn ($L_{hid} = 4$) với 100 nơ ron trên mỗi lớp, tổng số tham số Para là 50,600 và số phép tính FLOP là 100,000 . Trong khi đó, nếu tăng số lớp ẩn lên 10 ($L_{hid} = 10$), Para sẽ tăng lên 111,200 và FLOP đạt 220,000 . Ngoài ra, việc tăng số lớp ẩn cũng làm tăng đáng kể thời gian thực thi hệ thống, như đã minh họa trong Bảng 5.4. Do đó, để đạt được sự cân bằng giữa hiệu suất mô hình và chi phí tính toán, chúng ta lựa chọn số lớp ẩn $L_{hid} = 4$ làm số lớp ẩn tối ưu cho mô hình DNN. Hơn nữa, từ Hình 5.11, chúng ta quan sát được rằng đối với $BLER \approx 10^{-5}$, công suất được phân bổ khoảng 21dB . Điều này có nghĩa là công suất phát cho $\mathcal{U}_2$ và $\mathcal{U}_1$ lần lượt là 8.4 dB và 12.6 dB , tương ứng.

Bảng 5.4: Thời gian thực thi của mạng DNN, RMSE và MAPE theo số lớp ẩn khác nhau.

Trường hợp	Thời gian thực hiện (giây)	RMSE	MAPLE
DNN – 1 Lớp	0.007380	$1.805.10^{-8}$	$6.033.10^{-5}$
DNN – 2 Lớp	0.007694	$1.204.10^{-20}$	$4.389.10^{-16}$
DNN – 4 Lớp	0.01024	$1.204.10^{-20}$	$4.389.10^{-16}$
DNN – 10 Lớp	0.010244	$1.204.10^{-20}$	$4.389.10^{-16}$

5.6 Kết luận Chương 5

Chương 5 giới thiệu hai đóng góp đáng chú ý sau đây:

- Đóng góp thứ nhất là việc thiết kế một mô hình SPC mới có hỗ trợ của IRS, kết hợp một số kỹ thuật định dạng búp sóng tại nguồn phát và ứng dụng của kỹ thuật cải tiến NOMA. Mô hình hệ thống này được thiết kế để đáp ứng nhu cầu

của các mạng có giới hạn về công suất và hiệu quả sử dụng phổ, đồng thời đáp ứng nhu cầu của các hệ thống uRLLCs.

- Đóng góp thứ hai là đề xuất mô hình mạng DNN nhằm đáp ứng yêu cầu về tốc độ xử lý nhanh cho các mạng không dây thế hệ mới. Mạng DNN được thiết kế để giải quyết hai vấn đề cốt lõi như sau: (1) Dự đoán hiệu năng hệ thống; (2) Dự đoán công suất phân bổ nguồn phát. Kết quả nghiên cứu đã chứng minh rằng việc áp dụng mạng DNN không chỉ giảm đáng kể thời gian thực thi mà còn làm giảm độ phức tạp tính toán một cách hiệu quả.

Mô hình trình bày trong chương 5 phù hợp với các hệ thống mạng IoT và mạng thế hệ tương lai, nơi yêu cầu độ trễ thấp và chất lượng dịch vụ cao. Tuy nhiên, thực tế mô hình này gặp một số thách thức, đặc biệt trong việc thiết kế hệ thống đa anten để tạo chùm tia ở nguồn phát. Việc tối ưu chùm tia cần phải được cải thiện để đảm bảo hiệu quả hoạt động, ví dụ như tối ưu phân bổ tín hiệu và giảm nhiễu trong các môi trường phức tạp.

Nội dung của Chương 5 đã được công nhận thông qua bài báo được xuất bản trong tạp chí IET Communication, thuộc danh mục SCIE-Q2. Bài báo: “**N. T. Y. Linh**, P. N. Son, and V. N. Q. Bao, "*Intelligent reflecting surface-assisted beamforming-NOMA networks for short-packet communications: Performance analysis and deep learning approach*," vol. 17, no. 16, pp. 1940-1954, Aug. 2023, doi: 10.1049/cmu2.12667”.

Chương 6. KẾT LUẬN VÀ ĐỊNH HƯỚNG NGHIÊN CỨU TIẾP THEO

6.1 Giới thiệu

Chương 6 tổng kết những đóng góp khoa học Luận án và đề xuất các phương hướng nghiên cứu trong tương lai. Cụ thể, Mục 6.2 trình bày tóm tắt ngắn gọn về các nội dung đã đạt được và đồng thời nêu bật những đóng góp nổi bật của Luận án. Từ đó, Mục 6.3 tập trung vào việc định hướng mở rộng và phát triển Luận án trong tương lai.

6.2 Kết luận của Luận án

Luận án đã thành công trong việc đạt được hai mục tiêu khoa học theo đề cương được duyệt là:

- i) Xây dựng các mô hình SPC trong mạng truyền thông cộng tác và đề xuất phương pháp để tối ưu hiệu năng của hệ thống, nhằm đáp ứng yêu cầu của hệ thống uRLLCs.
- ii) Triển khai các kỹ thuật cải tiến hiệu năng ở lớp vật lý như mặt IRS, NOMA và DL, để tăng cường độ tin cậy và giảm thiểu độ trễ trong mạng thế hệ mới.

Về nội dung, Luận án đã đề xuất và tính toán thành công các chỉ số đặc trưng cho hiệu năng của ba mô hình SPC khác nhau trong mạng cộng tác. Cụ thể, đó là mô hình SPC trong mạng DF hai chặng, mạng chuyển tiếp qua IRS và NOMA. Đặc biệt, Luận án đã tập trung vào việc giải quyết thách thức về độ tin cậy và độ trễ trong SPC dưới tiêu chuẩn uRLLCs cho mạng 5G, thông qua các bước: 1) Khai thác phân tập không gian qua các giao thức tiềm năng như: TAS, PRS, MRC, IRS, và định dạng chùm tia; 2) Tối ưu công suất của nguồn phát và vị trí của nút chuyển tiếp để đảm bảo hiệu năng tốt nhất cho hệ thống; 3) Đề xuất kỹ thuật học sâu để thỏa mãn các yêu cầu về độ trễ thấp và tính chính xác cao trong hệ thống mạng không dây thế hệ 6G sắp tới.

Về phương pháp, Luận án đã thành công trong vận dụng nhiều phương pháp khác nhau để phân tích và đánh giá hiệu quả của ba mô hình cụ thể như sau: 1)

Đầu tiên, Luận án đã tiến hành so sánh, và đối chiếu các nghiên cứu liên quan đến SPC từ cả trong và ngoài nước. 2) Thứ hai, phương pháp toán giải tích, toán xác suất và tối ưu được áp dụng trong phân tích và tính hiệu năng cho các mô hình. Điều này hỗ trợ xây dựng lý thuyết cơ bản và cung cấp các chỉ số đánh giá hiệu năng chất lượng cho từng mô hình; 3) Cuối cùng, áp dụng mô phỏng thông qua phần mềm Matlab đã cho phép kiểm chứng lại các phân tích lý thuyết, từ đó xác nhận tính chính xác và ổn định của kết quả nghiên cứu, nâng cao độ tin cậy và sức thuyết phục của Luận án.

Về đóng góp, Luận án đã đem lại những cải tiến quan trọng trong giải quyết hai tiêu chí về độ tin cậy và thời gian trễ cho SPC, hướng đến đảm bảo yêu cầu uRLLCs như sau:

- ✓ Luận án đề xuất mô hình SPC trong mạng DF hai chặng trên kênh fading Rayleigh, mang lại các đóng góp như sau: Đầu tiên, Luận án chứng minh rằng hệ thống FDF cải thiện hiệu năng so với SDF với tất cả giá trị công suất được phân bổ của nguồn; Thứ hai, hiệu năng của hệ thống FDF cải thiện khi tăng số an-ten tại nguồn và đích, số nút chuyển tiếp, độ dài khối tin và công suất phát. Điểm nổi bật là hệ thống cung cấp độ lợi phân tập không gian, được tính bằng giá trị nhỏ nhất giữa tích số an-ten ở nguồn và số nút chuyển tiếp với số an-ten ở đích. Kết quả này cần thiết cho thiết kế hệ thống thực tế, giúp đảm bảo đạt được độ tin cậy và giảm chi phí triển khai hệ thống; Cuối cùng, Luận án cũng tìm được giá trị công suất của nguồn và vị trí của nút chuyển tiếp tối ưu cho trường hợp tổng quát. Kết quả này mang lại lợi ích lớn trong triển khai mạng thực tế, vì khi hệ thống được thiết lập ở chế độ tối ưu, hiệu năng đạt được là cao nhất và hoàn toàn đáp ứng các yêu cầu về độ tin cậy và độ trễ của dịch vụ uRLLCs trong hệ thống mạng thế hệ mới.
- ✓ Luận án đề xuất mô hình SPC trong mạng chuyển tiếp có hỗ trợ của IRS. Điểm mở rộng của hệ thống này đưa ra phương pháp để đánh giá hiệu năng, bao gồm BLER dạng chính xác và xấp xỉ, cùng với độ trễ của hệ thống trên kênh fading Rayleigh. Các kết quả chỉ ra rằng khi tăng công suất nguồn phát, các yếu tố phản xạ và chiều dài khối tin sẽ dẫn đến BLER giảm đáng kể. Đặc biệt, so sánh với FDF

và AF, hiệu năng hệ thống SPC qua IRS cải thiện đáng kể ở cùng công suất phát. Điều này mang nghĩa quan trọng trong thiết kế và triển khai mạng mới, cho phép thay thế nút chuyển tiếp bằng IRS để cải thiện QoS và mở rộng phạm vi phủ sóng với chi phí thấp hơn.

- ✓ Luận án đã thiết kế hệ thống SPC trong mạng NOMA có hỗ trợ của IRS. Điểm mới trong thiết kế của mô hình này là nguồn phát sử dụng kỹ thuật định dạng sóng nhằm tối ưu hiệu năng truyền tín hiệu đến các người dùng mong muốn. Hơn nữa, hệ thống cũng đã đề xuất các phương pháp để tính BLER, độ trễ và độ tin cậy. Đặc biệt, hệ thống đề xuất mạng DNN dựa trên DL để dự đoán hiệu năng và công suất phân bổ nguồn phát một cách chính xác, và đơn giản tính toán so với các phương pháp hiện có. Hơn nữa, nghiên cứu cũng chứng minh được rằng DL là cách tiếp cận hiệu quả, đáp ứng yêu cầu về thời gian thực trong mạng không dây hiện tại và tương lai. Kết luận này có ý nghĩa rất lớn trong việc phân bổ nguồn tài nguyên về công suất một cách thông minh và hiệu quả. Nói cách khác, dựa vào DL, chúng ta có thể tiên dự đoán các thông số như công suất, độ trễ hay thông lượng của hệ thống, tuân thủ các hạn chế về BLER theo yêu cầu uRLLCs, qua đó giảm thiểu chi phí khi thiết kế và triển khai mạng thực tiễn.

6.3 Hướng phát triển của Luận án

Luận án đã đạt được các mục tiêu đề ra. Tuy nhiên, trong bối cảnh nhu cầu QoS mạng của người dùng đang gia tăng, các công nghệ mạng vô tuyến ngày càng đa dạng hóa và phát triển. Do đó, trong thời gian tới NCS sẽ tiếp tục đề xuất các cải tiến cho hệ thống mạng đảm bảo tiêu chuẩn uRLLCs và cải tiến chất lượng cho các hệ thống mạng trong tương lai. Cụ thể, các nghiên cứu tiếp theo của Luận án sẽ dựa trên sự phát triển từ các hướng đã nghiên cứu của NCS như sau:

- Dựa vào công trình số 1, đó là: “*Intelligent Reflecting Surface-Assisted Beamforming-NOMA Networks for Short-Packet Communications: Performance Analysis and Deep Learning Approach*”, chúng ta có cơ sở để xây dựng mô hình phát theo các định hướng dưới đây:

- Mở rộng SPC trong mạng NOMA phục vụ đa người dùng phân bố ngẫu nhiên.
- Nghiên cứu mặt vừa phản xạ vừa truyền qua STAR-IRS thay thế IRS nhằm nâng cao vùng phủ sóng, tăng cường tối đa khả năng phục vụ đa người dùng trong hệ thống mạng. **Kết quả của hướng mở rộng này là công trình nghiên cứu số 6 trong danh mục công trình nghiên cứu của NCS.**
- Xem xét hệ thống SPC qua các kênh truyền fading tổng quát khác như kênh Nakagami-m và Rician.
- Nghiên cứu phương pháp học tăng cường nhằm giải quyết bài toán tối ưu hiệu năng hệ thống đáp ứng được dịch vụ uRLLCs cải tiến cho các hệ thống mạng tương lai như 6G. **Kết quả của hướng mở rộng này là công trình nghiên cứu số 7 trong danh mục công trình nghiên cứu của NCS.**
- Dựa vào công trình số 4, đó là: “*Phân Tích Độ Lợi Phân Tập Cho Mạng Chuyển Tiếp Qua Mặt Phản Xạ Thông Minh Và Nút Chuyển Tiếp Trong Truyền Thông Gói Tin Ngắn*” và công trình số 3 là: “*Dự Đoán Công Suất Tối Ưu Nguồn Phát Của Hệ Thống Truyền Gói Tin Ngắn Qua Mặt Phản Xạ Thông Minh Bằng Kỹ Thuật Học Sâu*”, mô hình này có thể mở rộng như sau:
 - Thiết kế hệ thống SPC được hỗ trợ bởi hệ nhiều tấm IRS nhằm hỗ trợ truyền tín hiệu ở nơi địa hình khó khăn như vùng đô thị lớn.
 - Xem xét giải quyết bài toán với điều kiện gần thực tế như CSI không hoàn hảo và qua kênh truyền Rician.
 - Giải quyết bài toán tối ưu hệ thống theo các tham số như góc dịch pha của hệ IRS, khoảng cách và công suất phân bổ nguồn của hệ thống.

Ngoài ra, nghiên cứu hệ thống SPC sử dụng sóng mili mét (mmWave) nhằm cải thiện về độ trễ và nâng cao tốc độ truyền dữ liệu cho các hệ thống mạng tương lai.

Hơn nữa, NCS xem xét phương pháp đa truy nhập phân chia tốc độ (RSMA) trong SPC nhằm tối ưu tốc độ truyền, hỗ trợ đa kết nối, giảm độ trễ và đặc biệt thực thi hệ thống mạng giảm phức tạp hơn phương pháp NOMA.

PHỤ LỤC

A. Chứng minh công thức (3.21)

Bằng cách đặt $\frac{P_S}{\sigma^2} = \alpha \frac{P_0}{\sigma^2} = \alpha \bar{\gamma}$, $\frac{P_{\mathcal{R}_b}}{\sigma^2} = (1-\alpha) \bar{\gamma}$ với $\alpha \in (0,1)$. Từ (3.17) và (3.20), bậc phân tập của hệ thống được biểu diễn dưới dạng như sau [89]:

$$\mathcal{DO} = -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\tilde{\mathcal{E}}_{e2e}^{(1)})_{(a)}}{\log(\bar{\gamma})} = -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log\left(\frac{\varsigma_1}{\bar{\gamma}^{NT}} + \frac{\varsigma_2}{\bar{\gamma}^M}\right)}{\log(\bar{\gamma})}, \quad (\text{A.1})$$

và

$$\mathcal{DO} = -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\tilde{\mathcal{E}}_{e2e}^{(2)})_{(a)}}{\log(\bar{\gamma})} = -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log\left(\frac{\varsigma_3}{\bar{\gamma}^{NT}} + \frac{\varsigma_4}{\bar{\gamma}^M}\right)}{\log(\bar{\gamma})}, \quad (\text{A.2})$$

với (a) là phép rút gọn bằng cách đặt như sau:

$$\varsigma_1 = \frac{v\sqrt{k}}{(NT+1)} \frac{(\rho_H^{NT+1} - \rho_L^{NT+1})}{(\Omega_{\mathcal{SR}}\alpha)^{NT}}. \quad (\text{A.3})$$

$$\varsigma_2 = \frac{v\sqrt{k}}{M!(M+1)} \frac{(\rho_H^{M+1} - \rho_L^{M+1})}{(\Omega_{\mathcal{R}_b\mathcal{D}}(1-\alpha))^M}. \quad (\text{A.4})$$

$$\varsigma_3 = \frac{\theta^{NT}}{(\Omega_{\mathcal{SR}}\alpha)^{NT}}. \quad (\text{A.5})$$

$$\varsigma_4 = \frac{\theta^M}{M!(\Omega_{\mathcal{R}_b\mathcal{D}}(1-\alpha))^M}. \quad (\text{A.6})$$

Bằng cách áp dụng biến đổi hàm logarit của (A.1), chúng ta được kết quả như sau:

$$\begin{aligned}
\mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^M + \varsigma_2 \bar{\gamma}^{NT}) - \log(\bar{\gamma}^{M+NT})}{\log(\bar{\gamma})}, \\
&= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^M + \varsigma_2 \bar{\gamma}^{NT})}{\log(\bar{\gamma})} - \frac{\log(\bar{\gamma}^{M+NT})}{\log(\bar{\gamma})}.
\end{aligned} \tag{A.7}$$

Chúng ta có 3 trường hợp có thể xảy ra cho (A.7) như sau:

- Trường hợp $NT = M$

$$\begin{aligned}
\mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^{NT} + \varsigma_2 \bar{\gamma}^{NT})}{\log(\bar{\gamma})} - \frac{\log(\bar{\gamma}^{NT+NT})}{\log(\bar{\gamma})}, \\
&= NT.
\end{aligned} \tag{A.8}$$

hay

$$\begin{aligned}
\mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^M + \varsigma_2 \bar{\gamma}^M)}{\log(\bar{\gamma})} - \frac{\log(\bar{\gamma}^{M+M})}{\log(\bar{\gamma})}, \\
&= M.
\end{aligned} \tag{A.9}$$

Từ (A.8) và (A.9), giá trị của $\mathcal{DO}$ của trường hợp $NT = M$ có thể tóm gọn lại là:

$$\mathcal{DO} = \min(NT, M). \tag{A.10}$$

- Trường hợp $NT > M$, bằng cách tính giới hạn $\bar{\gamma} \rightarrow \infty$ của (A.7), chúng ta thu được $\mathcal{DO}$ như:

$$\begin{aligned}
\mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^M + \varsigma_2 \bar{\gamma}^{NT})}{\log(\bar{\gamma})} - \frac{\log(\bar{\gamma}^{M+NT})}{\log(\bar{\gamma})} \\
&= -M + M + NT \\
&= NT.
\end{aligned} \tag{A.11}$$

- Trường hợp $NT < M$, tương tự cách tính như trong (A.11), chúng ta có:

$$\begin{aligned}
\mathcal{DO} &= -\lim_{\bar{\gamma} \rightarrow \infty} \frac{\log(\varsigma_1 \bar{\gamma}^M + \varsigma_2 \bar{\gamma}^{NT})}{\log(\bar{\gamma})} - \frac{\log(\bar{\gamma}^{M+NT})}{\log(\bar{\gamma})} \\
&= -NT + M + NT \\
&= M.
\end{aligned} \tag{A.12}$$

Bằng cách tổng hợp cho 3 trường hợp trong (A.10), (A.11), (A.12), chúng ta thu được kết quả như trong (3.21).

B. Chứng minh vấn đề tối ưu (3.24) là lời

Dựa trên [117, Định lý A.2], ma trận Hessian có dạng như sau:

$$\mathbf{H}(\mathbf{P}) = \begin{pmatrix} \frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_S^2} & \frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_S \partial P_{\mathcal{R}_b}} \\ \frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_{\mathcal{R}_b} \partial P_S} & \frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_{\mathcal{R}_b}^2} \end{pmatrix}, \quad (\text{B.1})$$

với

$$\frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_S^2} = \frac{NT(NT+1)(\theta\sigma^2 d_1^\eta)^{NT}}{P_S^{NT+2}} > 0, \quad (\text{B.2})$$

$$\frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_{\mathcal{R}_b}^2} = \frac{M(M+1)(\theta\sigma^2 d_1^\eta)^M}{M!P_{\mathcal{R}_b}^{M+2}} > 0, \quad (\text{B.3})$$

$$\frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_S \partial P_{\mathcal{R}_b}} = \frac{\partial^2 \tilde{\epsilon}_{e2e}(\mathbf{P})}{\partial P_{\mathcal{R}_b} \partial P_S} = 0. \quad (\text{B.4})$$

Chúng ta có thể thấy rõ ràng rằng các phần tử đường chéo của ma trận Hessian trong (B.1) là dương, thì ma trận $\mathbf{H}(\mathbf{P})$ dương. Do đó, hàm của $\tilde{\epsilon}_{e2e}(\mathbf{P})$ là hàm lồi.

C. Chứng minh công thức (3.29)

Dựa vào định nghĩa của đạo hàm, chúng ta có đạo hàm của hàm $f(\hat{x})$ theo biến $\hat{x}$ như sau:

$$\frac{\partial f(\hat{x})}{\partial(\hat{x})} = \lim_{x \rightarrow \hat{x}} \frac{f(x) - f(\hat{x})}{x - \hat{x}}. \quad (\text{C.1})$$

Điều này cũng có nghĩa là
$$\frac{\partial f(\hat{x})}{\partial(\hat{x})} \rightarrow \frac{f(x) - f(\hat{x})}{x - \hat{x}}, \quad (\text{C.2})$$

Ở đây ký hiệu $\rightarrow$ chỉ sự hội tụ. Giả định rằng $\hat{x}$ là một nghiệm của $f(\hat{x})$, Chúng ta có $f(\hat{x}) = 0$. Từ (C.2), chúng ta nhận được kết quả như sau:

$$x - \frac{f(x)}{\partial f(\hat{x}) / \partial \hat{x}} \rightarrow \hat{x}, \quad (\text{C.3})$$

trong đó $\hat{x}$ là nghiệm duy nhất hội tụ.

D. Chứng minh công thức (4.21)

Bằng cách thay thế (4.7) vào (4.24), e2e BLER của hệ thống SPC qua IRS có thể thu được bên dưới như sau:

$$\bar{\varepsilon}_{e2e} = v\sqrt{k} \int_{\rho_L}^{\rho_H} \left(1 - \frac{\Gamma\left(a+1, \frac{1}{b}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right)}{\Gamma(a+1)} \right) d\gamma. \quad (\text{D.1})$$

Bằng cách lấy tích phân, (D.1) được viết lại như sau:

$$\bar{\varepsilon}_{e2e}^{IRS} = v\sqrt{k}(\rho_H - \rho_L) - \frac{2v\sqrt{k}}{\Gamma(a+1)} \int_{\rho_L}^{\rho_H} \Gamma\left(a+1, \frac{1}{b}\sqrt{\frac{\gamma}{\bar{\gamma}}}\right) d\gamma, \quad (\text{D.2})$$

Tiếp theo, bằng cách đặt $y = \sqrt{\gamma}$ chúng ta đổi biến γ thành biến y và đặt

$$\mathcal{I} = \int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y \Gamma\left(a+1, \frac{y}{b\sqrt{\bar{\gamma}}}\right) dy \text{ trong (D.2) như sau:}$$

$$\bar{\varepsilon}_{e2e}^{IRS} = v\sqrt{k}(\rho_H - \rho_L) - \frac{2v\sqrt{k}}{\Gamma(a+1)} \underbrace{\int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y \Gamma\left(a+1, \frac{y}{b\sqrt{\bar{\gamma}}}\right) dy}_{\mathcal{I}}, \quad (\text{D.3})$$

Tiếp tục lấy tích phân từng phần trong (D.3), chúng ta đặt $u = \Gamma\left(a+1, \frac{y}{b\sqrt{\gamma}}\right)$ và

$dv = ydy$, tích phân $\mathcal{I} = uv - \int vdu$ được viết lại như sau:

$$\mathcal{I} = \frac{y^2}{2} \Gamma\left(a+1, \frac{y}{b\sqrt{\gamma}}\right) \Bigg|_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} - \frac{1}{2(b\sqrt{\gamma})^{a+1}} \underbrace{\int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y^{a+2} e^{-\frac{y}{b\sqrt{\gamma}}} dy}_{\mathcal{J}}, \quad (\text{D.4})$$

Bằng cách đặt $\mathcal{J} = \int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y^{a+2} e^{-\frac{y}{b\sqrt{\gamma}}} dy$ và sau đó, tích phân $\mathcal{J}$ được tính thông qua

phép lấy tích phân từng phần với $u = y^{a+2}$ và $dv = e^{-\frac{y}{b\sqrt{\gamma}}} dy$. Tích phân $\mathcal{J}$ được biến đổi như sau:

$$\mathcal{J} = -\frac{y^{a+2}}{2(b\sqrt{\gamma})^a} e^{-\frac{y}{b\sqrt{\gamma}}} \Bigg|_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} + \frac{a+2}{2(b\sqrt{\gamma})^a} \underbrace{\int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y^{a+1} e^{-\frac{y}{b\sqrt{\gamma}}} dy}_{\mathcal{J}^*}, \quad (\text{D.5})$$

Để ý thấy rằng, tích phân $\mathcal{J}^*$ trong (D.5) dựa vào Mathematica có giá trị như sau:

$$\mathcal{J}^* = \int_{\sqrt{\rho_L}}^{\sqrt{\rho_H}} y^{a+1} e^{-\frac{y}{b\sqrt{\gamma}}} dy = (b\sqrt{\gamma})^{2-a} \Gamma\left(a+2, \frac{y}{b\sqrt{\gamma}}\right) \Bigg|_{\sqrt{\rho_L}}^{\sqrt{\rho_H}}. \quad (\text{D.6})$$

Bằng cách thay thế (D.6), (D.5) vào (D.4), $\mathcal{I}$ trong (D.4) được tính như sau:

$$\mathcal{I} = \frac{1}{2} \left(y^2 \Gamma\left(a+1, \frac{y}{b\sqrt{\gamma}}\right) - \frac{y^{a+2}}{(b\sqrt{\gamma})^a} e^{-\frac{y}{b\sqrt{\gamma}}} - \frac{(a+2)}{\left(\frac{1}{b\sqrt{\gamma}}\right)^2} \Gamma\left(a+2, \frac{y}{b\sqrt{\gamma}}\right) \right) \Bigg|_{\sqrt{\rho_L}}^{\sqrt{\rho_H}}. \quad (\text{D.7})$$

Bằng cách đặt

$$\Delta(y) = \frac{1}{2} \left(y^2 \Gamma \left(a+1, \frac{y}{b\sqrt{\gamma}} \right) - \frac{y^{a+2}}{(b\sqrt{\gamma})^a} e^{\frac{-y}{b\sqrt{\gamma}}} - \frac{(a+2)}{\left(\frac{1}{b\sqrt{\gamma}} \right)^2} \Gamma \left(a+2, \frac{y}{b\sqrt{\gamma}} \right) \right). \quad (\text{D.8})$$

Bằng cách thay thế (D.8) và (D.7) vào (D.3), chúng ta thu được công thức như trong (4.25).

E. Chứng minh Bổ đề 1

Theo (5.15), BLER trung bình tại $\mathcal{U}_1$ có thể viết lại sau đây:

$$\bar{\varepsilon}_{\mathcal{U}_1} \approx v_1 \sqrt{k} \int_{\rho_{1L}}^{\rho_{1H}} F_{\gamma_{\mathcal{U}_1}^{\alpha_1}}(\gamma) d\gamma. \quad (\text{E.1})$$

Thay thế (5.10) vào (E.1), chúng ta thu được:

$$\bar{\varepsilon}_{\mathcal{U}_1}(x) = \begin{cases} v_1 \sqrt{k} \int_{\rho_{1L}}^{\rho_{1H}} dx, & \frac{\alpha_1}{\alpha_2} < \rho_{1L}, \\ v_1 \sqrt{k} \mathcal{I}_1 + v_1 \sqrt{k} \int_{\alpha_1/\alpha_2}^{\rho_{1H}} dx, & \rho_{1L} \leq \frac{\alpha_1}{\alpha_2} < \rho_{1H}, \\ v_1 \sqrt{k} \mathcal{I}_2, & \frac{\alpha_1}{\alpha_2} \geq \rho_{1H}, \end{cases} \quad (\text{E.2})$$

với

$$\mathcal{I}_1 = \int_{\rho_{1L}}^{\alpha_1/\alpha_2} \exp \left(-\lambda_0 \frac{x}{G_1 \bar{\gamma} (\alpha_1 - \alpha_2 x)} \right) dx, \quad (\text{E.3})$$

$$\text{và} \quad \mathcal{I}_2 = \int_{\rho_{1L}}^{\rho_{1H}} \exp \left(-\lambda_0 \frac{x}{G_1 \bar{\gamma} (\alpha_1 - \alpha_2 x)} \right) dx. \quad (\text{E.4})$$

Bằng cách đặt $y = \alpha_1 - \alpha_2 x$, chúng ta đổi biến x sang y . Tích phân (E.3) và (E.4) có thể được viết lại như sau:

$$\mathcal{I}_1 = \frac{1}{\alpha_2} \exp\left(\frac{\lambda_0}{\alpha_2 G_1 \bar{\gamma}}\right) \int_0^{\alpha_1 - \alpha_2 \rho_{1L}} \exp\left(-\lambda_0 \frac{\alpha_1}{G_1 \alpha_2 \bar{\gamma} y}\right) dy, \quad (\text{E.5})$$

$$\mathcal{I}_2 = \frac{1}{\alpha_2} \exp\left(\frac{\lambda_0}{\alpha_2 G_1 \bar{\gamma}}\right) \int_{\alpha_1 - \alpha_2 \rho_{1H}}^{\alpha_1 - \alpha_2 \rho_{1L}} \exp\left(-\lambda_0 \frac{\alpha_1}{G_1 \alpha_2 \bar{\gamma} y}\right) dy. \quad (\text{E.6})$$

Tiếp theo, bằng cách tính toán tích phân trong (E.5) và (E.6) sau đó thay thế $\mathcal{I}_1$ và $\mathcal{I}_2$ vào (E.2). Sau một số thao tác biến đổi, chúng ta có thể thu được **Bổ đề 1**.

F. Chứng minh Bổ đề 2

Dựa vào (5.11) và (5.12) và (5.15), BLER trung bình cho giải mã tín hiệu x_1 và x_2 tại $\mathcal{U}_2$ được dẫn ra như sau:

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_1}(x) = \begin{cases} v_2 \sqrt{k} (\rho_{2H} - \rho_{2L}), & \frac{\alpha_1}{\alpha_2} < \rho_{2L}, \\ v_2 \sqrt{k} \left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)} \mathcal{I}_3 \right], & \rho_{2L} \leq \frac{\alpha_1}{\alpha_2} < \rho_{2H}, \\ v_2 \sqrt{k} \left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)} \mathcal{I}_4 \right], & \frac{\alpha_1}{\alpha_2} \geq \rho_{2H}, \end{cases} \quad (\text{F.1})$$

với

$$\mathcal{I}_3 = v_2 \sqrt{k} \int_{\rho_{2L}}^{\frac{\alpha_1}{\alpha_2}} \Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{G_2 \bar{\gamma} (\alpha_1 - \alpha_2 x)}}\right) dx. \quad (\text{F.2})$$

$$\mathcal{I}_4 = v_2 \sqrt{k} \int_{\rho_{2L}}^{\rho_{2H}} \Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{G_2 \bar{\gamma} (\alpha_1 - \alpha_2 x)}}\right) dx. \quad (\text{F.3})$$

và

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_2} = v_2 \sqrt{k} \left((\rho_{2H} - \rho_{2L}) + \underbrace{\int_{\rho_{2L}}^{\rho_{2H}} \frac{\Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{\alpha_2 G_2 \bar{\gamma}}}\right)}{\Gamma(a+1)} dx}_{\mathcal{I}_5} \right). \quad (\text{F.4})$$

Bằng cách đặt $\mathcal{I}_5 = \int_{\rho_{2L}}^{\rho_{2H}} \frac{\Gamma\left(a+1, \frac{1}{b} \sqrt{\frac{x}{\alpha_2 G_2 \bar{\gamma}}}\right)}{\Gamma(a+1)} dx$, chúng ta có thể tính trực tiếp

trực tiếp $\mathcal{I}_5$ qua phần mềm Mathematica. Do đó, (F.4) được viết gọn lại sau đây:

$$\bar{\varepsilon}_{\mathcal{U}_2}^{x_2} = v_2 \sqrt{k} \left[(\rho_{2H} - \rho_{2L}) - \frac{2}{\Gamma(a+1)} \left(\Delta(\sqrt{\rho_{2H}}) - \Delta(\sqrt{\rho_{2L}}) \right) \right], \quad (\text{F.5})$$

với

$$\Delta(x) = \frac{1}{2} \left[x^2 \Gamma\left(a+1, \frac{x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right) - \frac{x^{a+2}}{\left(b\sqrt{\alpha_2 G_2 \bar{\gamma}}\right)^a} e^{\frac{-x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}} - \frac{(a+2)}{\left(\frac{1}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right)^2} \Gamma\left(a+2, \frac{x}{b\sqrt{\alpha_2 G_2 \bar{\gamma}}}\right) \right]. \quad (\text{F.6})$$

Tiếp theo, bằng cách đặt $z = \frac{1}{b} \sqrt{\frac{x}{G_2 \bar{\gamma} (\alpha_1 - \alpha_2 x)}}$, chúng ta có thể đổi cận tích

phân trong (F.2) và (F.3) như sau:

$$\mathcal{I}_3 = v_2 \sqrt{k} \int_{\mathfrak{I}_L}^{+\infty} \Gamma(a+1, z) \frac{2z}{\mathcal{Z} \alpha_2^2 (z^2 + \chi)^2} dz, \quad (\text{F.7})$$

$$\mathcal{I}_4 = v_2 \sqrt{k} \int_{\mathfrak{I}_L}^{\mathfrak{I}_H} \Gamma(a+1, z) \frac{2z}{\mathcal{Z} \alpha_2^2 (z^2 + \chi)^2} dz, \quad (\text{F.8})$$

với

$$\mathfrak{I}_L = \frac{1}{b} \sqrt{\frac{\rho_{2L}}{G_2 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{2L})}}. \quad (\text{F.9})$$

$$\mathfrak{I}_H = \frac{1}{b} \sqrt{\frac{\rho_{2H}}{G_2 \bar{\gamma} (\alpha_1 - \alpha_2 \rho_{2H})}}. \quad (\text{F.10})$$

$$\chi = \frac{1}{\mathcal{Z} \alpha_2}, \quad (\text{F.11})$$

$$\mathcal{Z} = G_2 \bar{\gamma} b^2. \quad (\text{F.12})$$

Hơn nữa, chúng ta áp dụng phương pháp tích phân từng phần bằng cách đặt

$u = \Gamma(a+1, z)$ và $dv = \frac{2z}{(z^2 + \chi)^2} dz$, (F.7) và (F.8) được phân tích chi tiết như sau:

$$\mathcal{I}_3 = \left(\frac{\alpha_1}{\mathcal{Z} \alpha_2^2} \right) \frac{\Gamma(a+1, \mathfrak{I}_L)}{\mathfrak{I}_L^2 + \chi} - \underbrace{\int_{\mathfrak{I}_L}^{+\infty} \frac{z^a e^{-z}}{z^2 + \chi} dz}_{\Psi_1}. \quad (\text{F.13})$$

$$\mathcal{I}_4 = \frac{\alpha_1}{\mathcal{Z} \alpha_2^2} \left[\frac{\Gamma(a+1, \mathfrak{I}_L)}{\mathfrak{I}_L^2 + \chi} - \frac{\Gamma(a+1, \mathfrak{I}_H)}{\mathfrak{I}_H^2 + \chi} \right] - \underbrace{\int_{\mathfrak{I}_L}^{\mathfrak{I}_H} \frac{z^a e^{-z}}{z^2 + \chi} dz}_{\Psi_2}. \quad (\text{F.14})$$

với $\Psi_1 = \int_{\mathfrak{I}_L}^{+\infty} \frac{z^a e^{-z}}{z^2 + \chi} dz$ và $\Psi_2 = \int_{\mathfrak{I}_L}^{\mathfrak{I}_H} \frac{z^a e^{-z}}{z^2 + \chi} dz$. Bằng cách đặt $u = e^{-z}$ và áp dụng tích

phân từng phần lần nữa, Ψ_1 và Ψ_2 được viết lại như sau:

$$\Psi_1 = \frac{1}{\chi(1+a)} \left(\left[z^{a+1} e^{-z} \right]_{\mathfrak{I}_L}^{\infty} + \int_{\mathfrak{I}_L}^{\infty} z^{a+1} e^{-z} \delta(z) dz \right), \quad (\text{F.15})$$

$$\Psi_2 = \frac{1}{\chi(1+a)} \left(\left[z^{a+1} e^{-z} \right]_{\mathfrak{I}_L}^{\mathfrak{I}_H} + \int_{\mathfrak{I}_L}^{\mathfrak{I}_H} z^{a+1} e^{-z} \delta(z) dz \right), \quad (\text{F.16})$$

với $\delta(z) = {}_2F_1\left(1, \frac{1+a}{2}; 1 + \frac{1+a}{2}; -\frac{z^2}{\chi}\right)$ với ${}_2F_1(.,.;.;.)$ là hàm Gauss

hypergeometric [85, CT (9.14.2)]. Theo [118], tích phân trong (F.15) và (F.16) là phân kỳ. Để tính các tích phân này, chúng ta sử dụng hàm int trong Matlab và tiếp theo, áp dụng hàm vpa để thu được kết quả xấp xỉ. Từ đó, chúng ta thu được giá trị xấp xỉ của Ψ_1 và Ψ_2 , tương ứng. Cuối cùng, bằng cách thay thế các kết quả xấp xỉ này vào (F.13) và (F.14), các kết quả thu được như trong **Bổ đề 2**.

TÀI LIỆU THAM KHẢO

- [1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE communications surveys & tutorials*, vol. 17, pp. 2347-2376, Jun. 2015, doi: 10.1109/COMST.2015.2444095.
- [2] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects," vol. 25, pp. 124-130, Jun. 2018, doi: 10.1109/MWC.2018.1700294.
- [3] M. Wang *et al.*, "Security and privacy in 6G networks: New areas and new challenges," vol. 6, pp. 281-291, Aug. 2020, doi: 10.1016/j.dcan.2020.07.003.
- [4] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, pp. 1617-1655, Aug. 2016, doi: 10.1109/COMST.2016.2532458.
- [5] J. Navarro-Ortiz *et al.*, "A survey on 5G usage scenarios and traffic models," vol. 22, pp. 905-929, Feb. 2020, doi: 10.1109/COMST.2020.2971781.
- [6] M. Carugi, "Key features and requirements of 5G/IMT-2020 networks," in *ITU Arab Forum on Emerging Technologies*, 2018, pp. 14-15: Internetsociety.org Reston, VA, USA.
- [7] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, pp. 1711-1726, Aug. 2016, doi: 10.1109/JPROC.2016.2537298.
- [8] H. Kim, "Ultra-Reliable and Low Latency Communication Systems," vol. pp. 303 - 342, 2020, doi: 10.1002/9781119494492.ch8.
- [9] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307, Apr. 2010, doi: 10.1109/TIT.2010.2043769.
- [10] R. Dong, C. She, W. Hardjawana, Y. Li, and B. J. I. T. o. W. C. Vucetic, "Deep learning for hybrid 5G services in mobile edge computing systems: Learn from a digital twin," vol. 18, pp. 4692-4707, Oct. 2019, doi: 10.1109/TWC.2019.2927312.
- [11] Y. Hu, J. Gross, and A. Schmeink, "On the performance advantage of relaying under the finite blocklength regime," *IEEE Communications Letters*, vol. 19, pp. 779-782, Mar. 2015, doi: 10.1109/LCOMM.2015.2410770.

- [12] P. Nouri, H. Alves, and M. Latva-aho, "On the performance of ultra-reliable decode and forward relaying under the finite blocklength," in *2017 European Conference on Networks and Communications (EuCNC)*, 2017, pp. 1-5: IEEE.
- [13] Y. Gu, H. Chen, Y. Li, and B. Vucetic, "Ultra-reliable short-packet communications: Half-duplex or full-duplex relaying?," *IEEE Wireless Communications Letters*, vol. 7, pp. 348-351, Jun. 2017, doi: 10.1109/LWC.2017.2777857.
- [14] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE communications Magazine*, vol. 42, pp. 74-80, Oct. 2004, doi: 10.1109/MCOM.2004.1341264.
- [15] J. N. Laneman, D. N. Tse, and G. W. J. I. T. o. I. t. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information theory*, vol. 50, pp. 3062-3080, Dec. 2004, doi: 10.1109/TIT.2004.838089.
- [16] M. Serror, C. Dombrowski, K. Wehrle, and J. Gross, "Channel coding versus cooperative ARQ: Reducing outage probability in ultra-low latency wireless communications," in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1-6: IEEE.
- [17] V. N. Swamy *et al.*, "Cooperative communication for high-reliability low-latency wireless control," in *2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 4380-4386: IEEE.
- [18] M. Serror, Y. Hu, C. Dombrowski, K. Wehrle, and J. Gross, "Performance analysis of cooperative ARQ systems for wireless industrial networks," in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2016, pp. 1-4: IEEE.
- [19] V. N. Q. Bao and T. T. Thanh, "Performance Analysis of Partial Relay Selection Networks with Short Packet Communications," *The 6th NAFOSTED Conference on Information and Computer Science (NICS)*, vol. pp. 23-26, Mar. 2019, doi: 10.1109/NICS48868.2019.9023894.
- [20] N. H. Tu and K. Lee, "Performance analysis and optimization of multihop MIMO relay networks in short-packet communications," *IEEE Transactions on Wireless Communications*, vol. pp. Dec. 2021, doi: 10.1109/TWC.2021.3131205.

- [21] K. Singh, M.-L. Ku, and M. F. Flanagan, "Resource Allocation in Energy-Efficient URLLC Multi-user Multicarrier AF Relay Networks," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1-6: IEEE.
- [22] Y. Hu, A. Schmeink, and J. Gross, "Blocklength-limited performance of relaying under quasi-static Rayleigh channels," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 4548-4558, Mar. 2016, doi: 10.1109/TWC.2016.2542245.
- [23] Y. Hu, A. Schmeink, and J. Gross, "Relaying with finite blocklength: Challenge vs. opportunity," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016, pp. 1-5: IEEE.
- [24] P. Nouri, H. Alves, R. D. Souza, and M. Latva-aho, "Ultra-reliable short message cooperative relaying protocols under Nakagami-m fading," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, 2017, pp. 287-292: IEEE.
- [25] B. Makki and M.-S. Alouini, "End-to-end Performance Analysis of Delay-sensitive Multi-relay Networks," *IEEE Communications Letters*, vol. 23, pp. 2159-2163, Dec. 2019, doi: 10.1109/LCOMM.2019.2945949.
- [26] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, pp. 106-112, Jan. 2019, doi: 10.1109/MCOM.001.1900107.
- [27] Q. Wu and R. J. I. t. o. w. c. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE transactions on wireless communications*, vol. 18, pp. 5394-5409, 2019, doi: 10.1109/TWC.2019.2936025.
- [28] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 2450-2525, Nov. 2020, doi: 10.1109/JSAC.2020.3007211.
- [29] N. S. Perović, M. Di Renzo, and M. F. Flanagan, "Channel capacity optimization using reconfigurable intelligent surfaces in indoor mmWave environments," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1-7: IEEE.

- [30] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753-116773, Aug. 2019, doi: 10.1109/ACCESS.2019.2935192.
- [31] E. Björnson, Ö. Ö, and E. G. Larsson, "Intelligent Reflecting Surface Versus Decode-and-Forward: How Large Surfaces are Needed to Beat Relaying?," *IEEE Wireless Communications Letters*, vol. 9, pp. 244-248, Feb. 2020, doi: 10.1109/LWC.2019.2950624.
- [32] M. Di Renzo *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," vol. 1, pp. 798-807, 2020, doi: 10.1109/OJCOMS.2020.3002955.
- [33] S. Zhou, W. Xu, K. Wang, M. Di Renzo, and M.-S. J. I. w. c. l. Alouini, "Spectral and energy efficiency of IRS-assisted MISO communication with hardware impairments," *IEEE wireless communications letters*, vol. 9, pp. 1366-1369, Sep. 2020, doi: 10.1109/LWC.2020.2990431.
- [34] H. Ren, K. Wang, and C. Pan, "Intelligent reflecting surface-aided URLLC in a factory automation scenario," *IEEE Transactions on Communications*, vol. 70, pp. 707-723, Jun. 2021, doi: 10.1109/TCOMM.2021.3125057.
- [35] T.-H. Vu, T.-V. Nguyen, D. B. da Costa, and S. Kim, "Intelligent reflecting surface-aided short-packet non-orthogonal multiple access systems," *IEEE Transactions on Vehicular Technology*, vol. 71, pp. 4500-4505, Apr. 2022, doi: 10.1109/TVT.2022.3146856.
- [36] T.-V. Nguyen, T.-N. Tran, K. Shim, T. Huynh-The, and B. An, "A Deep-Neural-Network-Based Relay Selection Scheme in Wireless-Powered Cognitive IoT Networks," *IEEE Internet of Things Journal*, vol. 8, pp. 7423-7436, May. 2020, doi: 10.1109/JIOT.2020.3038907.
- [37] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach," *IEEE Transactions on Wireless Communications*, vol. pp. Jul. 2021, doi: 10.1109/TWC.2021.3060514.
- [38] C. Sun and C. Yang, "Learning to optimize with unsupervised learning: Training deep neural networks for URLLC," in *2019 IEEE 30th Annual*

- International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1-7: IEEE.
- [39] C. Sun and C. Yang, "Unsupervised deep learning for ultra-reliable and low-latency communications," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1-6: IEEE.
 - [40] M. Y. Abdelsadek, Y. Gadallah, and M. H. Ahmed, "Resource Allocation of URLLC and eMBB Mixed Traffic in 5G Networks: A Deep Learning Approach," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*, 2020, pp. 1-6: IEEE.
 - [41] H. V. Hóa, N. H. Tú, and V. N. Q. Bảo, "Đánh Giá Hiệu Năng Mạng Chuyển Tiếp Đa Chặng Trong Truyền Thông Gói Tin Ngắn," vol. 1, pp. 43-49, Dec. 2020, doi.
 - [42] P. M. Quang and T. T. Duy, "Nâng Cao Hiệu Năng Mạng Chuyển Tiếp Đa Chặng Truyền Thông Gói Tin Ngắn Với Truyền Thông Cộng Tác Tăng Cường Tại Mỗi Chặng," 2023, vol. 1, no. 1, pp. 66-73.
 - [43] P. N. Son, T. T. Duy, P. V. Tuan, and T. P. J. E. J. Huynh, "Short packet communication in underlay cognitive network assisted by an intelligent reflecting surface," *ETRI Journal*, vol. 45, pp. 28-44, Sep. 2023, doi: 10.4218/etrij.2021-0435.
 - [44] V. N. Q. B. Huỳnh Văn Hóa, "Ước lượng hiệu năng bằng kỹ thuật học sâu cho hệ thống mặt phản xạ thông minh RIS kết hợp thiết bị bay không người lái UAV trong truyền thông gói tin ngắn," in *Journal of Science and Technology on Information and Communications*, 2022, vol. 1, pp. 57-66.
 - [45] T.-H. Vu, T.-V. Nguyen, T.-T. Nguyen, and S. Kim, "Performance Analysis and Deep Learning Design of Wireless Powered Cognitive NOMA IoT Short-Packet Communications with Imperfect CSI and SIC," *IEEE Internet of Things Journal*, vol. 9, pp. Jul. 2021, doi: 10.1109/JIOT.2021.3121421.
 - [46] T.-V. Nguyen, V.-D. Nguyen, D. B. da Costa, and B. An, "Short-Packet Communications in Multi-Hop WPINs: Performance Analysis and Deep Learning Design," 2021 *IEEE Global Communications Conference (GLOBECOM)*, vol. pp. 1-6, Feb.
Doi: 10.1109/GLOBECOM46510.2021.9685765.
 - [47] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE*

- Transactions on Communications, vol. 64, pp. 618-629, Dec. 2015, doi: 10.1109/TCOMM.2015.2511087.
- [48] P. C. Sofotasios and S. Freear, "Novel expressions for the Marcum and one dimensional Q-functions," in *2010 7th International Symposium on Wireless Communication Systems*, 2010, pp. 736-740: IEEE.
 - [49] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Transactions on Information Theory*, vol. 61, pp. 2430-2438, May 2015, doi: 10.1109/TIT.2015.2411256.
 - [50] M. C. Gursoy, "Throughput analysis of buffer-constrained wireless systems in the finite blocklength regime," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, pp. 1-13, Dec. 2013, doi: 10.1186/1687-1499-2013-290.
 - [51] B. Makki, T. Svensson, and M. Zorzi, "Finite block-length analysis of the incremental redundancy HARQ," *IEEE Wireless Communications Letters*, vol. 3, pp. 529-532, Aug. 2014, doi: 10.1109/LWC.2014.2353059.
 - [52] A. Adinoyi and H. Yanikomeroglu, "Cooperative relaying in multi-antenna fixed relay networks," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 533-544, 2007, doi: 10.1109/TWC.2007.05227.
 - [53] B.-K. An, T. T. Duy, and H.-Y. Kong, "A cooperative transmission strategy using entropy-based relay selection in mobile ad-hoc wireless sensor networks with Rayleigh fading environments," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 3, pp. 147-162, Apr. 2009, doi: 10.3837/tiis.2009.02.002
 - [54] M. Serror, S. Vaaßen, K. Wehrle, and J. Gross, "Practical evaluation of cooperative communication for ultra-reliability and low-latency," in *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*, 2018, pp. 14-15: IEEE.
 - [55] Y. Lu and W. H. Mow, "Low-complexity detection and performance analysis for decode-and-forward relay networks," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4819-4823: IEEE.
 - [56] T. Q. Duong, D. B. da Costa, M. El Kashlan, and V. N. Q. Bao, "Cognitive amplify-and-forward relay networks over Nakagami- m fading," *IEEE*

- Transactions on Vehicular Technology, vol. 61, pp. 2368-2374, Jun. 2012, doi: 10.1109/TVT.2012.2192509.
- [57] T. L. Thanh, V. N. Q. Bao, and T. T. Duy, "Capacity analysis of multi-hop decode-and-forward over Rician fading channels," in *2014 international conference on computing, management and telecommunications (ComManTel)*, 2014, pp. 134-139: IEEE.
 - [58] T. L. Thanh, T. Duy, V. Bao, and N. Nhat, "Outage performance with selection decode and forward in cognitive radio with imperfect csi," in *Proc. of ICGHIT2014*, 2014, pp. 156-161.
 - [59] M. K. Simon and M.-S. J. I. T. o. I. T. Alouini, *Digital communications over fading channels* (no. 7). 2008, pp. 3369-3370.
 - [60] G. J. Foschini and M. J. J. W. p. c. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless personal communications*, 1998, vol. 6, pp. 311-335.
 - [61] M. M. Al-Wani, A. Sali, N. K. Noordin, S. J. Hashim, C. Y. Leow, and I. J. I. A. Krikidis, "Robust beamforming and user clustering for guaranteed fairness in downlink NOMA with partial feedback," vol. 7, pp. 121599-121611, Aug. 2019, doi: 10.1109/ACCESS.2019.2936911.
 - [62] Y. Zhao, R. Adve, and T. J. J. I. c. I. Lim, "Outage probability at arbitrary SNR with cooperative diversity," vol. 9, pp. 700-702, Aug. 2005, doi: 10.1109/LCOMM.2005.1496587.
 - [63] J. Hu and N. C. J. I. C. L. Beaulieu, "Performance analysis of decode-and-forward relaying with selection combining," vol. 11, pp. 489-491, Jul. 2007, doi: 10.1109/LCOMM.2007.070065.
 - [64] A. Bletsas, A. Khisti, D. P. Reed, and A. J. I. J. o. s. a. i. c. Lippman, "A simple cooperative diversity method based on network path selection," vol. 24, pp. 659-672, Mar. 2006, doi: 10.1109/JSAC.2005.862417.
 - [65] V. N. Q. Bao and H. Y. Kong, "Performance analysis of multi-hop decode-and-forward relaying with selection combining," *Journal of Communications and Networks*, vol. 12, pp. 616-623, Dec. 2010, doi: 10.1109/JCN.2010.6388309.
 - [66] V. N. Q. Bao and H. Y. Kong, "Diversity order analysis of dual-hop relaying with partial relay selection," *IEICE transactions on communications*, vol. 92, pp. 3942-3946, Dec. 2009, doi: 10.1587/transcom.E92.B.3942.

- [67] T. T. Duy and V. N. Q. Bao, "Multi-hop transmission with diversity combining techniques under interference constraint," in *2013 international conference on advanced technologies for communications (ATC 2013)*, 2013, pp. 131-135: IEEE.
- [68] T. T. Duy and H.-Y. Kong, "Cooperative Communication with Different Combining Techniques in One-Dimensional Random Networks," *Journal of electromagnetic engineering and science*, vol. 12, pp. 13-19, Mar. 2012, doi: 10.5515/JKIEES.2012.12.1.13
- [69] B. Q. V. Nguyen and H. Y. J. I. t. o. c. Kong, "BER performance of decode-and-forward relaying using equal-gain combining over rayleigh fading channels," *IEICE transactions on communications*, vol. 91, pp. 3760-3763, Nov. 2008, doi: 10.1093/ietcom/e91-b.11.3760.
- [70] H. Yang *et al.*, "Design of resistor-loaded reflectarray elements for both amplitude and phase control," vol. 16, pp. 1159-1162, Nov. 2016, doi: 10.1109/LAWP.2016.2626318.
- [71] W. Tang *et al.*, "Wireless Communications with Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement," *IEEE Transactions on Wireless Communications*, vol. pp. 1-1, Sep. 2020, doi: 10.1109/TWC.2020.3024887.
- [72] R. Alghamdi *et al.*, "Intelligent surfaces for 6G wireless networks: A survey of optimization and performance analysis techniques," vol. 8, pp. 202795-202818, Oct. 2020, doi: 10.1109/ACCESS.2020.3031959.
- [73] H. Zhang, B. Di, Z. Han, H. V. Poor, and L. J. I. w. c. l. Song, "Reconfigurable intelligent surface assisted multi-user communications: How many reflective elements do we need?," vol. 10, pp. 1098-1102, Feb. 2021, doi: 10.1109/LWC.2021.3058637.
- [74] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE transactions on communications*, vol. 69, pp. 3313-3351, 2021, doi: 10.1109/TCOMM.2021.3051897.
- [75] Y. Liu *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE communications surveys & tutorials*, vol. 23, pp. 1546-1577, May. 2020, doi: 10.1109/COMST.2021.3077737.

- [76] B. N. Raju, B. N. Rao, and M. Basha, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 12, pp. 8145-8191, Oct. 2021, doi.
- [77] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Communications Surveys & Tutorials* vol. 20, pp. 2294-2323, May. 2018, doi: 10.1109/COMST.2018.2835558.
- [78] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, pp. 2595-2621, Jun. 2018, doi: 10.1109/COMST.2018.2846401.
- [79] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," in *nature* 2015, vol. 521, no. 7553, pp. 436-444.
- [80] Y. Gu, H. Chen, Y. Li, L. Song, and B. Vucetic, "Short-packet two-way amplify-and-forward relaying," *IEEE Signal Processing Letters*, vol. 25, pp. 263-267, 2017, doi.
- [81] V. N. Q. Bảo, T. T. J. J. o. S. Thanh, T. o. Information, and Communications, "Mạng Chuyển Tiếp Đa Chặng Dạng Nền Trong Truyền Thông Gói Tin Ngắn: Đánh Giá Tỷ Lệ Lỗi Khó," vol. 1, pp. 69-74, 2020, doi.
- [82] N. T. Y. Linh, N. H. Tú, V. N. Q. J. J. o. S. Bảo, T. o. Information, and Communications, "Đánh Giá Hiệu Năng Mạng Chuyển Tiếp Từng Phần Với Đa Anten Thu Trong Truyền Thông Gói Tin Ngắn," in *Journal of Science and Technology on Information and Communications*, 2020, vol. 1, no. 4A, pp. 53-61.
- [83] I. Krikidis, J. Thompson, S. McLaughlin, and N. Goertz, "Amplify-and-forward with partial relay selection," *IEEE Communications letters*, vol. 12, pp. 235-237, Apr. 2008, doi: 10.1109/LCOMM.2008.071987.
- [84] W. Yang, G. Durisi, T. Koch, and Y. J. I. T. o. I. T. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," vol. 60, pp. 4232-4265, Jul. 2014, doi: 10.1109/TIT.2014.2318726.
- [85] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2014.
- [86] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

- [87] S. Huang, B. Sheng, and C. Ji, "Channel Coding at Finite Blocklength and Its Application in 6G URLLC," *Journal of Computer Communications*, vol. 11, pp. 132-142, Sep. 2023, doi: 10.1109/TIT.2010.2043769.
- [88] B. Makki, T. Svensson, G. Caire, and M. Zorzi, "Fast HARQ over finite blocklength codes: A technique for low-latency reliable communication," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 194-209, 2018, doi.
- [89] D. Gündüz, M. A. Khojastepour, A. Goldsmith, and H. V. J. I. T. o. W. C. Poor, "Multi-hop MIMO relay networks: diversity-multiplexing trade-off analysis," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1738-1747, May. 2010, doi: 10.1109/TWC.2010.05.090915.
- [90] P. Popovski *et al.*, "Wireless access for ultra-reliable low-latency communication: Principles and building blocks," *Ieee Network*, vol. 32, pp. 16-23, Apr. 2018, doi: 10.1109/MNET.2018.1700258.
- [91] S. S. Sastry, *Introductory methods of numerical analysis*. PHI Learning Pvt. Ltd., 2012.
- [92] J. Ehiwario and S. J. I. J. o. E. Aghamie, "Comparative study of bisection, Newton-Raphson and secant methods of root-finding problems," in *IOSR Journal of Engineering*, 2014, vol. 4, no. 4, pp. 01-07.
- [93] A. Goldsmith, *Wireless communications* (Copyright by Cambridge University Press). Stanford University, 2005.
- [94] O. L. A. López, E. M. G. Fernández, R. D. Souza, and H. Alves, "Ultra-reliable cooperative short-packet communications with wireless energy transfer," *IEEE Sensors Journal*, vol. 18, pp. 2161-2177, Mar. 2018, doi: 10.1109/JSEN.2018.2789480.
- [95] L. Yang, Y. Yang, M. O. Hasna, and M.-S. Alouini, "Coverage, Probability of SNR Gain, and DOR Analysis of RIS-Aided Communication Systems," *IEEE Wireless Communications Letters*, vol. pp. 2020, doi.
- [96] A.-A. A. Boulogeorgos and A. Alexiou, "Performance Analysis of Reconfigurable Intelligent Surface-Assisted Wireless Systems and Comparison With Relaying," *IEEE Access*, vol. 8, pp. 94463-94483, May. 2020, doi: 10.1109/ACCESS.2020.2995435.

- [97] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing indoor spectrum sharing capacity using smart reflect-array," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1-6: IEEE.
- [98] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Communications Magazine*, vol. 58, pp. 106-112, 2020, doi: 10.1109/MCOM.001.1900107.
- [99] T. Sharma, A. Chehri, and P. J. E. Fortier, "Reconfigurable intelligent surfaces for 5G and beyond wireless communications: A comprehensive survey," vol. 14, pp. 8219, Dec. 2021, doi: 10.3390/en14248219.
- [100] A. Papoulis and H. Saunders, *Probability, random variables and stochastic processes*. 1989.
- [101] S. Primak, V. Kontorovich, and V. Lyandres, "Stochastic methods and their applications to communications," in *Stochastic Differential Equations Approach*. John Wiley&Sons, 2004.
- [102] T. T. Duy, G. C. Alexandropoulos, V. T. Tung, V. N. Son, and T. Q. J. I. C. Duong, "Outage performance of cognitive cooperative networks with relay selection over double-Rayleigh fading channels," *IET Communications*, vol. 10, pp. 57-64, Jan 2016, doi: 10.1049/iet-com.2015.0236.
- [103] M. O. Hasna and M.-S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh-fading channels," *IEEE transactions on Wireless Communications*, vol. 2, pp. 1126-1131, Nov 2003, doi: 10.1109/TWC.2003.819030.
- [104] V. N. Q. Bao, *Mô phỏng hệ thống truyền thông*. Hà Nội: Nhà xuất bản khoa học và kỹ thuật, p. 267.
- [105] T. S. Rappaport, *Wireless communications: principles and practice*, 2 ed. Cambridge University Press, 2024.
- [106] C. D. Ho, T.-V. Nguyen, T. Huynh-The, T.-T. Nguyen, D. B. da Costa, and B. An, "Short-packet communications in wireless-powered cognitive IoT networks: Performance analysis and deep learning evaluation," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 2894-2899, Feb. 2021, doi: 10.1109/TVT.2021.3061157.
- [107] C. A. Balanis, *Antenna theory: analysis and design*. John wiley & sons, 2015.

- [108] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [109] J. M. Steele, *The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities*. Cambridge University Press, 2004.
- [110] H. Chen *et al.*, "Ultra-reliable low latency cellular networks: Use cases, challenges and approaches," *IEEE Communications Magazine*, vol. 56, pp. 119-125, Dec. 2018, doi: 10.1109/MCOM.2018.1701178.
- [111] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," in *towards data science*, 2017, vol. 6, no. 12, pp. 310-316.
- [112] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications surveys & tutorials*, vol. 21, pp. 2224-2287, Mar. 2019, doi: 10.1109/COMST.2019.2904897.
- [113] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, pp. e623, 2021, doi: doi.org/10.7717/peerj-cs.623.
- [114] B. Zargar, F. Ponci, A. J. I. T. o. I. Monti, and Measurement, "Evaluation of computational complexity for distribution systems state estimation," vol. 72, pp. 1-12, Mar. 2023, doi: 10.1109/TIM.2023.3260268.
- [115] N. T. Y. Linh, T. Ngo Hoang, P. N. Son, and V. N. Q. Bao, "Dual-hop Relaying Networks for Short-Packet URLLCs: Performance analysis and Optimization," *Journal of Communications and Networks*, vol. 24, pp. 408 - 418, Jul. 2022, doi: 10.23919/JCN.2022.000020.
- [116] C. She *et al.*, "A Tutorial on Ultrareliable and Low-Latency Communications in 6G: Integrating Domain Knowledge Into Deep Learning," in *Proceedings of the IEEE*, 2021, vol. 109, no. 3, pp. 204-246.
- [117] S. S. Rao, *Engineering optimization: theory and practice*. John Wiley & Sons, 2019.
- [118] M. Abramowitz, "Handbook of mathematical functions with formulas," in *Graphs, and Mathematical Tables*, ed, 1965.

DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA NCS

Tạp chí Quốc tế

- [1]. **N. T. Y. Linh**, P. N. Son, and V. N. Q. Bao, "*Intelligent reflecting surface-assisted beamforming-NOMA networks for short-packet communications: Performance analysis and deep learning approach*," IET Communications, vol. 17, no. 16, pp. 1940-1954, Aug.2023, doi: 10.1049/cmu2.12667. **(SCIE-Q2)**
- [2]. **N. T. Y. Linh**, T. Ngo Hoang, P. N. Son, and V. N. Q. Bao, "*Dual-hop Relaying Networks for Short-Packet URLLCs: Performance analysis and Optimization*," Journal of Communications and Networks, vol. 24, no. 4, pp. 408-418, Aug. 2022, doi: 10.23919/JCN.2022.000020. **(SCIE-Q1)**

Tạp chí trong nước

- [3]. **N. T. Y. Linh**, V. N. Q. Bảo, and P. N. Son, "*Dự đoán công suất tối ưu nguồn phát của hệ thống truyền gói tin ngắn qua mặt phản xạ thông minh bằng kỹ thuật học sâu*," Journal of Science and Technology on Information and Communications (JSTIC), vol. 1, no. 1A, pp. 11-18, Apr. 2023.
- [4]. **N. T. Y. Linh**, V. N. Q. Bao, Pham Ngoc Son, "*Phân Tích Độ Lợi Phân Tập Cho Mạng Chuyển Tiếp Qua Mặt Phản Xạ Thông Minh Và Nút Chuyển Tiếp Trong Truyền Thông Gói Tin Ngắn*," Journal of Science and Technology on Information and Communications (JSTIC), vol. 1, no. 4B, pp. 56-64, Dec.2021.
- [5]. **N. T. Y. Linh**, N. H. Tu, and V. N. Q. Bảo, "*Đánh Giá Hiệu Năng Mạng Chuyển Tiếp Từng Phần Với Đa An-ten Thu Trong Truyền Thông Gói Tin Ngắn*," Journal of Science and Technology on Information and Communications (JSTIC), vol. 1, no. 4A, pp. 53-61, Dec.2020.

Kỷ yếu hội nghị quốc tế

- [6]. **N. T. Y. Linh**, P. N. Son, and V. N. Q. Bao, "*Performance Evaluation of Simultaneously Transmitting and Reflecting Intelligent Reflecting Surface - Assisted NOMA Networks for Short-Packet Communications*," in 2023 RIVF

International Conference on Computing and Communication Technologies (RIVF), 2023, pp. 194-199, doi: 10.1109/RIVF60135.2023.10471771.

- [7]. **N. T. Y. Linh**, P. N. Son, and V. N. Q. Bao, "*Power allocation for the STAR-RIS-NOMA Short Packet communication system using Deep Reinforcement Learning*," in 2024 International Conference on Advanced Technologies for Communications (ATC), 2024, pp. 215-220, IEEE.