

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

PERSPECTIVES PRATIQUES SUR LE FONCTIONNEMENT DES RIS PASSIFS
AUTONOMES: APPROCHE DE ADVERSARIAL BANDIT ASSISTÉ PAR
ALGORITHME GÉNÉTIQUE

MÉMOIRE PRÉSENTÉ(E)
COMME EXIGENCE PARTIELLE DE MAÎTRISE EN GÉNIE ÉLECTRIQUE

PAR
NADA BELHADJ LTAIEF

AVRIL 2025

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire, de cette thèse ou de cet essai a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire, de sa thèse ou de son essai.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire, cette thèse ou cet essai. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire, de cette thèse et de son essai requiert son autorisation.

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

MAÎTRISE EN GÉNIE ÉLECTRIQUE (M. Sc. A.)

Direction de recherche :

Messaoud Ahmed Ouameur	Directeur de recherche, UQTR
------------------------	------------------------------

Daniel Massicotte	Co-directeur de recherche, UQTR
-------------------	---------------------------------

Jury d'évaluation

Daniel Massicotte	Co-directeur de recherche, UQTR
-------------------	---------------------------------

Frédéric Domingue	Évaluateur interne, UQTR
-------------------	--------------------------

Sébastien Roy	Évaluateur externe, Université de Sherbrook
---------------	---------------------------------------------

Résumé

Les communications sans fil jouent un rôle central dans les réseaux de nouvelle génération, notamment avec l'avènement des technologies 5G et 6G. Dans ce contexte, les surfaces intelligentes reconfigurables (RIS) émergent comme une solution révolutionnaire pour améliorer les performances des réseaux. Ces dispositifs novateurs permettent de manipuler les signaux radio pour renforcer la couverture, réduire les interférences et maximiser l'efficacité des transmissions.

Cependant, l'intégration des RIS dans des environnements réels soulève des défis complexes liés à leur contrôle et à l'optimisation de leurs configurations. Avant toute optimisation, il est essentiel de garantir le bon fonctionnement des RIS, car les grandes tailles des surfaces et les nombreuses possibilités de réglage des éléments de phase créent un espace de solution immense. Cela rend les approches traditionnelles inefficaces ou trop coûteuses en termes de calcul.

L'objectif principal de cette étude est de développer et d'évaluer des algorithmes avancés capables d'optimiser efficacement les configurations des RIS pour maximiser la puissance reçue, tout en assurant leur fonctionnement stable et fiable dans des environnements réels.

Pour relever ce défi, des expériences réelles ont été réalisées. Ces expériences ont utilisé un système intégrant des RIS contrôlés et testés avant d'entamer l'optimisation. Plusieurs techniques d'optimisation ont été mises en œuvre et comparées, notamment les Algorithmes Génétiques (GA), les réseaux de neurones profonds avec apprentissage par renforcement (DQN), les Bandits Adversaires (AB), ainsi qu'une approche hybride innovante, les Bandits Adversaires assistés par GA (GA-aided AB). Ces méthodes ont été évaluées dans des environnements expérimentaux simulant des scénarios réels, avec

différentes positions d'émetteurs et de récepteurs en termes d'angles et de distances. Les performances des algorithmes ont été analysées selon leur capacité à la force du signal et à converger rapidement vers des configurations optimales.

Parmi les algorithmes testés, l'approche hybride AB assisté par le GA a démontré des performances supérieures. En combinant l'exploration approfondie des GA et la robustesse des Bandits Adversaires, cette méthode s'est avérée efficace pour naviguer dans l'espace de solution complexe et identifier des configurations optimales, même dans des conditions difficiles.

En conclusion, cette étude met en lumière le rôle crucial des RIS et des algorithmes hybrides dans la transformation des réseaux de communication sans fil. En s'appuyant sur des expérimentations réelles, elle souligne l'importance d'explorer des techniques avancées pour relever les défis technologiques et tirer pleinement parti des possibilités offertes par cette technologie émergente.

Remerciement

Je tiens à exprimer toute ma gratitude à celles et ceux qui m'ont soutenue et encouragée tout au long de l'élaboration de ce mémoire.

Avant tout, je remercie chaleureusement mon père Slaheddine dont le soutien moral, la patience et la confiance en mes capacités m'ont profondément inspirée. Je remercie également ma mère Fouzia pour son amour inconditionnel, sa bienveillance et ses précieux conseils qui m'ont accompagnée à chaque étape de ce parcours.

Un grand merci à mes sœurs Rahma et Lina, qui ont su me motiver et me réconforter dans les moments de doute, et dont l'affection a été un pilier indispensable à ma réussite.

Je souhaite aussi exprimer ma sincère reconnaissance à professeur Ahmed Messaoud Ouameur, qui a su me guider avec compétence et disponibilité. Son expertise et ses remarques constructives ont largement contribué à l'avancement et à la qualité de ce travail.

Enfin, j'adresse ma gratitude à toutes les personnes qui, de près ou de loin, ont participé à cette aventure et m'ont permis de franchir cette étape importante de ma formation.

Table des matières

Résumé	iii
Avant Propos	v
Table des matières	vi
Liste des tableaux	x
Liste des figures	xi
Liste des symboles	xiii
Chapitre 1 - Introduction	1
1.1 Contexte général et problématique	1
1.2 Problématique	2
1.3 Objectifs de recherche	3
1.4 Méthodologie de recherche	4
1.5 Structure du mémoire	5
1.6 Conclusion	6
Chapitre 2 - Revue de littérature	7
2.1 Introduction	7
2.2 Les surfaces intelligentes reconfigurables	7
2.2.1 Architecture du RIS	8
2.2.1.1 Classification des RIS.....	9
2.2.1.2 RIS passif ou actif	10

2.2.1.3	Conception et contrôle géométriques	10
2.2.1.4	Adaptabilité en temps réel	11
2.2.2	Modes de fonctionnement du RIS	12
2.2.2.1	Région de champ proche.....	12
2.2.2.2	Région du champ lointain	12
2.2.2.3	Beamforming:	13
2.2.2.4	Réflexion:	14
2.2.2.5	Réfraction:	14
2.2.2.6	Absorption:	14
2.2.3	Principe de fonctionnement du RIS	15
2.2.3.1	Concept de base de fonctionnement	15
2.2.3.2	Contrôle du déphasage.....	15
2.2.4	Les applications du RIS	18
2.3	Travaux antérieurs	20
2.4	Contribution	22
2.5	Contexte institutionnel et continuité des travaux	24
2.6	Conclusion	24
Chapitre 3 - Méthodologie		26
3.1	Introduction	26
3.2	Cadre du système et formulation du problème	26
3.3	Ressources matérielles et logiciels pour le fonctionnement du système	28
3.3.1	Éléments constitutifs du système	29
3.3.1.1	MAX7301: Extenseur GPIO.....	29
3.3.1.2	Module Software-Defined Radio	30
3.3.1.3	Antenne unidirectionnel	31

3.3.1.4	Microcontrôleur	31
3.3.2	Protocoles utilisés pour l'interconnexion	32
3.3.2.1	TCP/IP	32
3.3.2.2	MQTT	32
3.3.2.3	SPI	33
3.3.3	Langages de Programmation et Outils Logiciels	33
3.4	Conception et contrôle RIS	34
3.4.1	RIS personnalisé pour la réalisation du projet	34
3.4.2	Couches matérielles du RIS	35
3.4.3	Contrôle de déphasage basé sur une diode PIN	36
3.4.4	Contrôle et configuration du RIS	37
3.4.5	Le contrôle du RIS via MAX7301	38
3.5	Principe du fonctionnement du système	39
3.6	Stratégies d'Optimisation pour le RIS	41
3.6.1	GA comme approche de base	41
3.6.2	DRL et DRL assistés par GA comme méthodes comparatives	46
3.6.2.1	Préliminaires du DRL	46
3.6.2.2	DQN pour l'optimisation RIS: une description algorithmique	50
3.6.2.3	DQN assisté par GA pour l'optimisation RIS: une description algorithmique	54
3.6.3	Bandit Adversaire	56
3.6.4	AB assisté par GA	60
3.7	Conclusion	64
Chapitre 4 - Expérimentation et résultats		65

4.1	Originalité de l'approche expérimentale	65
4.2	Description de l'installation expérimentale	65
4.2.1	Positionnement des équipements	66
4.2.2	Génération du signal	67
4.2.3	Réception et optimisation	68
4.3	Comparaison des performances des méthodes d'optimisation RIS dans les configurations spatiales	68
4.3.1	Performances générales des algorithmes d'optimisation RIS	72
4.3.2	Impact des positions spatiales sur les performances des méthodes d'optimisation RIS	73
4.4	Comportement de convergence des méthodes d'optimisation assistées par GA et autonomes	75
4.5	Synthèse comparative des performances des algorithmes	77
4.6	Conclusion	77
Chapitre 5 - Conclusion		81
Annexe A - La connexion SPI Daisy chain et RIS		88

Liste des tableaux

Tableau 4-1 Positions du récepteur par rapport au RIS	67
Tableau 4-2 Tableau des paramètres pour les algorithmes GA, DQN et EXP3 ...	70
Tableau 4-3 Comparaison des performances des algorithmes d'optimisation RIS avec valeurs minimales et maximales des gains	72
Tableau 4-4 Résumé des Performances des Algorithmes d'Optimisation RIS	80

Liste des figures

Figure 2.1	Les modes d'opération du RIS	13
Figure 2.2	Illustration du mécanisme de réflexion du RIS et de la configuration des déphasages	16
Figure 2.3	Les différents applications du RIS	18
Figure 3.1	Les pins du MAX7301	29
Figure 3.2	Configuration SPI avec maître et esclave	33
Figure 3.3	RIS personnalisé	34
Figure 3.4	Les dimensions du RIS et des UC	35
Figure 3.5	Architecture matérielle du RIS	36
Figure 3.6	Description du fonctionnement de l'installation expérimentale	40
Figure 3.7	Architecture du RIS assistée par GA	42
Figure 3.8	Architecture du RIS assistée par DQN	53
Figure 3.9	Architecture du DQN assistée par GA	55
Figure 3.10	Architecture du AB assistée par GA	61
Figure 4.1	Vue de dessus de l'installation	66
Figure 4.2	Vue latérale de l'installation	66
Figure 4.3	Configuration avec différents points de vue	66
Figure 4.4	Positions de test de Pos1 à Pos6	69
Figure 4.5	Diagramme en barres comparant les valeurs CMA RSSI pour diverses méthodes d'optimisation à plusieurs positions de récepteur	71
Figure 4.6	Performance du gain CMA des méthodes d'optimisation RIS à différentes positions	71

Figure 4.7	Courbes de convergence des algorithmes d'optimisation pour six positions	79
------------	--------------------------------------------------------------------------------	----

Liste des symboles

AB	Adversarial Bandit (Bandit Adversarial)
ARIS	Absorbent RIS (RIS absorbant)
BER	Bit Error Rate (Taux d'Erreur Binaire)
BS	Base Station (Station de Base)
CMA	Cumulative Moving Average (Moyenne Cumulative Mobile)
DNN	Deep Neural Network (Réseau Neuronal Profond)
DQN	Deep Q-Network (Réseau Q Profond)
DRL	Deep Reinforcement Learning (Apprentissage par Renforcement Profond)
EM	Électromagnétique
EXP3	Exponential-weight algorithm for Exploration and Exploitation (Algorithme Pondéré Exponentiel pour l'Exploration et l'Exploitation)
GA	Genetic Algorithm (Algorithme Génétique)
G-MAQ	Genetic Multi-Agent Q-learning (Apprentissage Q Multi-Agent Génétique)
GPIO	General Purpose Input Output (Entrée-Sortie Polyvalente)
IP	Internet Protocol (Protocole Internet)
IoT	Internet of Things (Internet des Objets)
LED	Light-Emitting Diode (Diode Électroluminescente)
LoS	Line-of-Sight (Ligne de Visée direct)
LR	Learning Rate (Taux d'Apprentissage)
M2M	Machine-to-Machine (Machine à Machine)
MAB	Multi-Armed Bandit (Bandit Multi-Bras)
MEMS	Micro-Electro-Mechanical Systems (Systèmes Micro-Électro-Mécaniques)
MEC	Mobile Edge Computing (Calcul en Périphérie Mobile)

MIMO Multiple-Input Multiple-Output (Entrées Multiples, Sorties Multiples)

ML Machine Learning (Apprentissage automatique)

mmWave Millimeter Wave (Onde Millimétrique)

MQTT Message Queuing Telemetry Transport (Protocole de Transport de Télémétrie en File d'Attente)

NGA Niching Genetic Algorithm (Algorithme Génétique de Nichage)

NOMA Non-Orthogonal Multiple Access (Accès Multiple Non-Orthogonal)

PIN Positive-Intrinsic-Negative (Positif-Intrinsèque-Négatif)

PLS Physical Layer Security (Sécurité de la Couche Physique)

PSO Particle Swarm Optimization (Optimisation par Essaim de Particules)

ReLU Rectified Linear Unit (Unité Linéaire Rectifiée)

RF Radio Frequency (Fréquence Radio)

RIS Reconfigurable Intelligent Surface (Surface Intelligente Reconfigurable)

RLC Résistance-Inductance-Capacité

RL Reinforcement Learning (Apprentissage par Renforcement)

RSSI Received Signal Strength Indicator (Indicateur de la Force du Signal Reçu)

SDR Software-Defined Radio (Radio Définie par Logiciel)

SNR Signal-to-Noise Ratio (Rapport Signal-Bruit)

SPI Serial Peripheral Interface (Interface Périphérique Série)

SRE Smart Radio Environment (Environnement Radio Intelligent)

SWIPT Simultaneous Wireless Information and Power Transfer (Transmission Simultanée d'Information et d'Énergie Sans Fil)

TCP Transmission Control Protocol (Protocole de Contrôle de Transmission)

THz Terahertz

UAV Unmanned Aerial Vehicle (Véhicule Aérien Sans Pilote)

UC unit cell (cellule unitaire)

V2X Vehicle-to-Everything (Véhicule-à-Tout)

Chapitre 1 - Introduction

1.1 Contexte général et problématique

Depuis la 2G jusqu'à la 4G, chaque génération de réseaux mobiles a apporté des avancées majeures, comme la communication vocale et la messagerie (2G), l'Internet mobile (3G) et les applications gourmandes en données telles que le streaming vidéo et les jeux en ligne (4G) [1]. Cependant, ces réseaux montrent leurs limites face à la croissance exponentielle des appareils connectés et aux exigences croissantes en termes de débit et de capacité [2].

La 5G, déployée depuis 2019, a introduit des améliorations significatives, notamment des vitesses de communication élevées, une faible latence et une connectivité élargie [3]. Cependant, elle reste confrontée à des défis tels que la couverture restreinte, les blocages de signaux et les coûts d'infrastructure élevés. Ces limitations ont conduit à une exploration intensive des réseaux au-delà de la 5G (B5G) et de la 6G, qui visent à offrir une connectivité omniprésente, des débits ultra-élevés et une gestion intelligente basée sur l'IA [4–6]. Ces réseaux promettent également une meilleure durabilité et efficacité énergétique pour répondre aux défis environnementaux croissants.

Dans ce contexte, les RIS (Reconfigurable Intelligent Surface) émergent comme une solution révolutionnaire pour redéfinir la communication sans fil [7]. En manipulant dynamiquement les ondes EM (électromagnétique) à travers des éléments contrôlables, les RIS améliorent la force du signal, réduisent les interférences et optimisent l'efficacité énergétique [8]. Leur faible consommation d'énergie, leur adaptabilité et leur intégration transparente dans l'infrastructure existante leur permettent de surmonter des obstacles tels que les blocages de signaux et les zones de couverture limitée. Grâce à ces caractéristiques, les RIS s'imposent comme une technologie clé pour les réseaux de communication

de nouvelle génération, offrant une couverture étendue, une capacité améliorée et une efficacité énergétique accrue.

Cette étude met ainsi l'accent sur l'intégration et l'optimisation des RIS, en explorant leur rôle fondamental dans la réalisation des objectifs ambitieux des futurs systèmes de communication, tout en répondant aux défis liés à la performance, aux coûts et à l'impact environnemental.

1.2 Problématique

Bien que les RIS présentent un potentiel unique pour surmonter les défis des réseaux B5G et 6G, leur optimisation reste un problème complexe. L'espace de configuration des RIS, qui croît de manière exponentielle avec le nombre d'éléments qui le constituent, complique considérablement le développement de stratégies efficaces. De plus, la nécessité d'une réactivité en temps réel dans des environnements sans fil dynamiques ajoute une couche de difficulté, en particulier lorsqu'il s'agit de garantir à la fois l'efficacité informatique et l'évolutivité.

Les approches existantes rencontrent souvent des difficultés à concilier la rapidité de calcul et l'exploration approfondie de l'espace de configuration. Cela limite leur capacité à s'adapter aux conditions environnementales changeantes et aux exigences diverses des applications modernes. La nature dynamique des systèmes sans fil, marquée par des évolutions rapides des canaux et des besoins variables en termes de performance, aggrave encore ces défis.

Cette étude vise à relever ces défis en explorant de nouvelles méthodologies d'optimisation adaptées aux besoins spécifiques des RIS. L'objectif est de maximiser la force du signal, d'améliorer l'efficacité énergétique et de garantir une adaptabilité en temps

réel tout en maintenant des exigences informatiques gérables. Une question clé se pose, comment concevoir et mettre en œuvre des algorithmes capables d'optimiser efficacement les configurations RIS en temps réel, tout en répondant aux exigences de performance et d'évolutivité ?

En s'appuyant sur des techniques d'optimisation avancées, cette recherche cherche à proposer des solutions robustes et innovantes pour exploiter pleinement le potentiel des RIS dans les réseaux de communication de nouvelle génération

1.3 Objectifs de recherche

L'objectif principal de ce projet est de concevoir, développer et évaluer des approches algorithmiques avancées pour optimiser la configuration des RIS, dans le but de maximiser les performances des systèmes de communication sans fil. Cette optimisation vise à améliorer des paramètres clés tels que la force du signal reçu, l'efficacité énergétique et la gestion des interférences, tout en garantissant une adaptabilité en temps réel aux environnements sans fil dynamiques. Pour atteindre cet objectif, le projet s'articule autour de plusieurs axes :

- Identifier les défis spécifiques liés à la configuration des RIS, notamment le vaste espace de configuration et les exigences de réactivité en temps réel.
- Développer des algorithmes avancés, intégrant des techniques heuristiques, évolutives et d'apprentissage automatique, capables d'explorer et d'exploiter efficacement l'espace de configuration des RIS.
- Valider expérimentalement les approches proposées dans des scénarios contrôlés.
- Évaluer l'impact des solutions développées sur des métriques de performance spécifiques, telles que le débit de données, la couverture réseau, la robustesse aux blocages de signaux et l'efficacité globale des systèmes.

Ce projet ambitionne non seulement de proposer des solutions robustes et innovantes pour l'optimisation des RIS, mais aussi de fournir des recommandations pratiques pour leur déploiement et leur intégration dans des réseaux sans fil complexes.

1.4 Méthodologie de recherche

La méthodologie adoptée dans ce mémoire repose sur une approche structurée et progressive pour répondre à la problématique et atteindre les objectifs fixés. Elle se décline en plusieurs étapes clés :

. Analyse de la problématique et revue de la littérature

Une analyse approfondie de la problématique a été réalisée pour identifier les défis spécifiques liés à l'optimisation des RIS dans les réseaux sans fil. Cette étape inclut une revue de la littérature exhaustive visant à comprendre le principe de fonctionnement des RIS, leurs applications, et les méthodes d'optimisation existantes. Cela a permis de situer le projet dans le contexte des technologies de nouvelle génération (5G et 6G) et de mettre en évidence les lacunes dans les travaux antérieurs.

. Mise en place du contrôle des RIS et de la plateforme expérimentale

Le contrôle du RIS utilisé dans le projet a été implémenté, intégrant une plateforme expérimentale adaptée. Cette étape comprenait la configuration et le déploiement de la surface reconfigurable ainsi que l'intégration des composants nécessaires pour permettre un ajustement dynamique des configurations. Cela a permis de créer un environnement expérimental contrôlé, essentiel pour évaluer la faisabilité et les performances des algorithmes proposés.

. Conception et implémentation des algorithmes d'optimisation

Des algorithmes d'optimisation avancés ont été développés pour répondre aux besoins spécifiques des RIS. Ces algorithmes, intégrant des techniques heuristiques et des

approches inspirées de l'apprentissage automatique, ont été directement implémentés dans l'environnement expérimental.

. Validation et déploiement des algorithmes dans l'environnement expérimental

Les algorithmes développés ont été déployés dans l'environnement expérimental, en tenant compte des contraintes spécifiques des réseaux sans fil dynamiques. Cette étape a permis de valider leur robustesse, leur applicabilité et leurs performances dans des conditions réelles, tout en respectant les exigences liées à l'adaptabilité et au temps réel.

. Collecte et analyse des résultats expérimentaux

Enfin, des campagnes expérimentales ont été menées pour collecter des données sur les performances des algorithmes dans des conditions contrôlées. Les résultats obtenus ont été analysés pour évaluer l'impact des approches proposées sur des métriques clés, telles que la force du signal, l'efficacité énergétique et la robustesse aux interférences. Ces analyses ont servi à valider les contributions du projet et à formuler des recommandations pour des applications futures.

1.5 Structure du mémoire

La structure de ce mémoire est organisée de manière logique pour guider le lecteur à travers les différentes étapes de la recherche. Le premier chapitre, introduction, établit le contexte général de l'étude, a exposé la problématique et les objectifs de recherche, et présente une vue d'ensemble des défis liés à l'optimisation des RIS. Le deuxième chapitre est la revue de littérature, explore les travaux existants sur les réseaux sans fil assistés par RIS, en mettant en lumière les avancées technologiques, les méthodes d'optimisation et les lacunes identifiées, tout en justifiant la contribution spécifique de cette recherche. Le troisième chapitre, réalisation du projet, détaille la méthodologie adoptée, y compris la conception et l'implémentation des algorithmes, ainsi que la mise en place de la plateforme

expérimentale utilisée pour tester et valider ces approches. Le quatrième chapitre, résultats de tests, présente et analyse les données expérimentales collectées, en évaluant l'efficacité et les performances des algorithmes développés selon des métriques pertinentes. Enfin, la Conclusion récapitule les principales contributions de l'étude, discute des implications pratiques et scientifiques, et propose des perspectives pour des recherches futures dans le domaine de l'optimisation des RIS. Cette structure permet une progression claire et cohérente du problème initial vers les solutions proposées et leurs validations.

1.6 Conclusion

Ce chapitre a présenté les enjeux des réseaux sans fil de nouvelle génération et le rôle prometteur des RIS pour surmonter leurs limites. Il a mis en lumière les défis d'optimisation associés et défini les objectifs de recherche pour concevoir des solutions innovantes. Les chapitres suivants approfondiront l'état de l'art, les réalisations et les résultats expérimentaux de cette étude.

Chapitre 2 - Revue de littérature

2.1 Introduction

Ce chapitre examine en détail les concepts essentiels et les éléments pertinents tirés de la littérature qui doivent être pris en considération lors du contrôle et l'optimisation du RIS. Il discute le RIS en détail, en nous concentrant sur son architecture matérielle, les différents modes de fonctionnement et les systèmes RIS proposés et fabriqués. Ces éléments sont essentiels pour comprendre le fonctionnement pratique des RIS et leur rôle dans les systèmes de communication sans fil modernes.

2.2 Les surfaces intelligentes reconfigurables

Les RIS ont suscité un vif intérêt au sein de la communauté des communications sans fil en raison de leur capacité à configurer des SRE (Smart Radio Environment), où les conditions de propagation peuvent être adaptées pour optimiser la communication [9]. Ce concept introduit une nouvelle vision de l'environnement sans fil, permettant de personnaliser le canal de propagation grâce au déploiement du RIS. En façonnant le front d'onde des ondes radio à travers le réseau, les RIS offrent la possibilité de reconfigurer l'environnement sans fil afin de répondre aux exigences spécifiques du système. Selon [9], un SRE est défini comme un environnement radio reconfigurable et intelligent, capable de jouer un rôle actif dans le transfert et le traitement des informations. Cette approche améliore la fiabilité de l'échange de données entre les émetteurs et les récepteurs, en transformant ainsi la manière dont les communications sans fil sont envisagées. Les principaux avantages du RIS sont les suivants:

- **Configuration rentable et économe en énergie:** RIS permet de configurer l'environnement de propagation sans fil de manière rentable et avec une faible consommation d'énergie, ce qui en fait une technologie innovante pour le

développement futur des réseaux sans fil. En utilisant une conception passive, les RIS éliminent la nécessité d'une électronique complexe, facilitant ainsi le déploiement et réduisant les dépenses du système.

- **Intégration transparente:** La conception fine et légère des RIS permet une intégration transparente dans diverses infrastructures, allant des bâtiments et des tours de télécommunications aux systèmes satellites, aux installations de villes intelligentes et aux réseaux de transport.
- **Installation polyvalente:** Les RIS peuvent être installés stratégiquement sur de grandes surfaces planes, que ce soit à l'intérieur ou à l'extérieur, sur des murs, des plafonds ou des bâtiments entiers, pour refléter l'énergie du signal sans fil de manière stratégique et surmonter les obstacles, créant ainsi un chemin de propagation virtuel simulant une LoS (Line-of-Sight) [10].
- **Programmabilité et ajustements dynamiques:** La possibilité de programmer les RIS améliore leur utilité en permettant des ajustements en temps réel pour l'optimisation sans fil. Cela garantit une configuration simple et une extensibilité, permettant au RIS d'être flexible dans divers contextes.
- **Gestion à distance:** Certaines solutions RIS offrent des capacités de gestion à distance, améliorant l'accessibilité et simplifiant la maintenance [11].

2.2.1 Architecture du RIS

Les RIS agissent comme des structures programmables bidimensionnelles (2D) capables de modifier leur réponse aux ondes EM en temps réel. Cette adaptabilité unique distingue les réseaux assistés par RIS des systèmes de communication traditionnels en permettant un contrôle intelligent des canaux entre émetteurs et récepteurs. Grâce à cette technologie, il est possible d'améliorer la force du signal reçu sur les terminaux en

manipulant les réflexions du signal et en appliquant des techniques avancées de formation de faisceaux [12], [13].

2.2.1.1 *Classification des RIS*

Les architectures matérielles RIS sont principalement classées en fonction de la composition matérielle et de la conception structurelle de leurs éléments réfléchissants, ce qui conduit à deux catégories principales, les structures basées sur un réseau d'antennes [14] et les structures basées sur des métasurfaces [15].

Structures basées sur un réseau d'antennes : Ces architectures reposent sur des réseaux compacts d'antennes capables de contrôler les réflexions des signaux en ajustant leur phase et leur amplitude. L'espacement entre les antennes, optimisé à une demi-longueur d'onde, réduit les interférences et améliore la focalisation des faisceaux réfléchis, permettant une communication directionnelle efficace. Ces structures sont particulièrement adaptées aux bandes du mmWave (Millimeter Wave) et THz (Terahertz), où les petites longueurs d'onde permettent des réseaux d'antennes très denses [16].

Structures basées sur la métasurface : Les métasurfaces, formées de métamatériaux, permettent de contrôler les ondes EM avec une précision exceptionnelle grâce à leurs éléments de taille inférieure à la longueur d'onde. Elles peuvent plier, focaliser et remodeler les ondes d'une manière qu'il est impossible de reproduire avec des matériaux conventionnels. Cela les rend particulièrement utiles pour manipuler finement les signaux dans des environnements complexes, comme les zones urbaines denses, et pour des applications nécessitant une orientation précise et une mise en forme avancée des faisceaux [15, 16].

2.2.1.2 *RIS passif ou actif*

Une autre distinction cruciale dans les architectures RIS réside dans le fait que le système soit passif ou actif. Cette classification a un impact significatif sur les performances, l'efficacité énergétique et la complexité du système.

RIS passif : Les RIS passifs réfléchissent les signaux sans amplification, ajustant uniquement leur phase et direction. Économes en énergie, ils ne nécessitent pas d'alimentation externe, mais souffrent de pertes dues à l'*effet de double trajet*, ce qui limite leur efficacité sur de longues distances [13, 17].

RIS actif : Les RIS actifs, équipés d'amplificateurs, compensent les pertes en amplifiant les signaux réfléchis. Ils offrent de meilleures performances sur de longues distances ou dans des environnements complexes, mais leur consommation d'énergie est plus élevée en raison des composants actifs [17, 18].

2.2.1.3 *Conception et contrôle géométriques*

La réponse du signal des éléments RIS, telle que leur amplitude de réflexion et leur déphasage, peut être contrôlée avec précision grâce à des variations dans la géométrie des éléments. La forme, la taille, la disposition et l'orientation des éléments individuels du réseau RIS jouent un rôle central dans la détermination de la manière dont le système manipule les ondes EM incidentes [19]. Cette capacité permet au RIS d'effectuer une formation de faisceau dynamique, où la direction et la focalisation du signal réfléchi sont ajustées en permanence pour optimiser les performances de communication.

Dans les systèmes de communication sans fil, le *coefficient de réflexion* de chaque élément, défini comme le rapport entre le signal réfléchi et le signal incident, détermine la façon dont le RIS interagit avec les ondes EM entrantes. En ajustant les coefficients de réflexion,

RIS peut optimiser la qualité du signal à l'extrémité du récepteur, en dirigeant le signal vers l'emplacement souhaité tout en minimisant les interférences et la dégradation du signal [16]. Ces ajustements peuvent être effectués en temps réel, permettant au RIS de s'adapter aux changements de l'environnement sans fil, tels que la mobilité des utilisateurs, la variabilité des canaux et les obstructions environnementales.

2.2.1.4 *Adaptabilité en temps réel*

L'une des caractéristiques les plus intéressantes du RIS est sa capacité à s'adapter en temps réel aux changements de l'environnement sans fil. À mesure que les utilisateurs se déplacent ou que les conditions environnementales changent, les canaux sans fil, par lesquels les signaux se propagent, changent également de manière dynamique. Cela peut entraîner des fluctuations dans la force, la qualité et la couverture du signal. Pour maintenir des performances de communication optimales, le RIS doit être capable d'ajuster ses coefficients de réflexion en temps réel, affinant les signaux réfléchis pour tenir compte de ces variations [13].

Atteindre cette adaptabilité en temps réel nécessite l'intégration de composants électroniques avancés, tels que des *diodes PIN (Positive-Intrinsic-Negative)*, des *varactors* ou des *MEMS (Micro-Electro-Mechanical Systems)*, qui permettent une commutation rapide et un réglage continu des coefficients de réflexion. Ces composants permettent à chaque élément du RIS de modifier dynamiquement son déphasage et son amplitude, offrant ainsi des capacités de formation de faisceaux en temps réel qui améliorent les performances globales du système de communication [16]. Cette adaptabilité est particulièrement importante dans les scénarios mobiles, où la position des utilisateurs change fréquemment, nécessitant des ajustements continus des chemins de signal.

2.2.2 Modes de fonctionnement du RIS

La technologie RIS fonctionne selon plusieurs modes clés, chacun adapté à une fonction spécifique de manipulation des ondes EM. Les principaux modes de fonctionnement du RIS. Lorsque l'on discute des modes de fonctionnement du RIS, il est essentiel de comprendre la distinction physique entre les régions en champ proche et en champ lointain.

2.2.2.1 Région de champ proche

La propagation du signal des RIS se divise en régions de champ proche et de champ lointain, différenciées par la distance de Rayleigh, définie comme $d_{\text{Rayleigh}} = \frac{2D^2}{\lambda}$, où D est l'ouverture du RIS et λ la longueur d'onde.

Un canal émetteur-RIS-récepteur est en champ proche si la distance émetteur-RIS ou RIS-récepteur est inférieure à la distance de Rayleigh [20]. Dans les anciens réseaux sans fil, le champ proche avait un impact limité en raison des fréquences basses et des antennes plus petites, rendant la propagation en champ lointain dominante, avec des ondes planes favorisant la simplicité des conceptions [21]. Avec l'avènement des RIS pour les réseaux 6G, les ouvertures plus grandes offrent des gains significatifs, notamment pour les bandes mmWave et THz [22], qui réduisent le blocage des signaux. Ainsi, le champ proche jouera un rôle majeur dans les futurs systèmes, avec une portée élargie et des performances améliorées.

2.2.2.2 Région du champ lointain

La région du champ lointain est la zone autour d'une antenne ou d'un RIS où les ondes EM se propagent sous forme d'ondes planes, avec des fronts d'onde plats. Située

au-delà de la distance de Rayleigh, proportionnelle au carré de l'antenne ou de l'ouverture RIS et inversement proportionnelle à la longueur d'onde, elle offre une propagation prévisible avec une phase uniforme, essentielle pour les conceptions traditionnelles de communications sans fil [22]. Prisée dans les systèmes conventionnels pour la simplicité de la modélisation des canaux et l'orientation efficace des faisceaux, elle reste centrale même avec les réseaux 6G. Cependant, avec des antennes et RIS plus grands, le champ proche doit aussi être pris en compte pour maximiser les performances, notamment à des fréquences plus élevées [21].

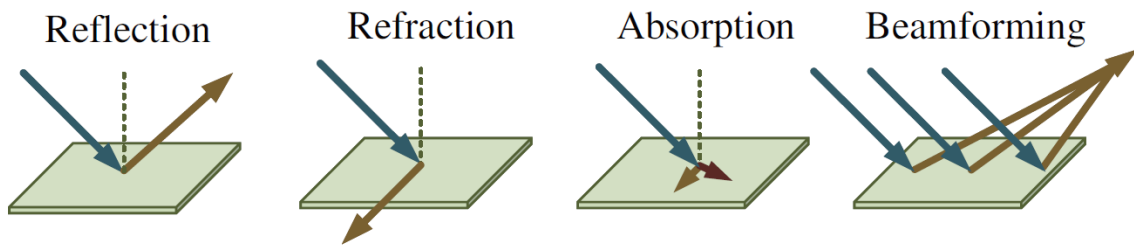


Figure 2.1 Les modes d'opération du RIS

2.2.2.3 *Beamforming*:

La formation de faisceau permet au RIS de concentrer les ondes réfléchies vers un point spécifique, améliorant ainsi la réception du signal. Chaque UC applique des déphasages uniques, calculés en fonction de la géométrie du RIS et de la position cible, pour provoquer une interférence constructive au niveau souhaité. Cette approche optimise la force du signal et la qualité de la communication. Essentielle pour les communications point à point et multi-utilisateurs, cette technique offre une directivité et une puissance accrues [23–25]. Elle est particulièrement pertinente pour les réseaux 6G, où un contrôle précis des ondes est nécessaire pour maximiser les performances et gérer les interférences.

2.2.2.4 *Réflexion:*

La réflexion dans RIS fait référence à la redirection des ondes incidentes en ajustant la phase de chaque élément sur la surface RIS. Ce processus permet à l'onde d'être réfléchi dans une direction spécifique, qui peut être contrôlée dynamiquement. Contrairement aux réflexions traditionnelles où l'angle de réflexion est déterminé par l'angle d'incidence, RIS peut orienter le faisceau dans une direction souhaitée en appliquant différents déphasages à chaque élément, réalisant ainsi la formation de faisceau ou l'orientation du faisceau. Un exemple pratique de ceci peut être vu dans [26], où le coefficient de réflexion est modulé dynamiquement pour obtenir un contrôle précis des fronts d'onde et de la manipulation harmonique.

2.2.2.5 *Réfraction:*

La réfraction dans RIS implique la transmission d'ondes EM à travers la surface avec un changement de direction en raison de la manipulation de phase à travers les éléments. Semblable à la réfraction traditionnelle, le RIS permet aux ondes transmises de se plier selon différents angles, mais ici, l'angle et la direction peuvent être contrôlés pour optimiser la propagation du signal. Un exemple de ceci est démontré dans [27] qui explique comment les métasurfaces peuvent obtenir un contrôle précis sur le comportement des ondes réfractées.

2.2.2.6 *Absorption:*

L'absorption dans les RIS désigne leur capacité à absorber une partie des ondes incidentes, plutôt que de les réfléchir entièrement. En ajustant la phase et l'amplitude des signaux, le RIS peut réduire les interférences et optimiser l'efficacité spectrale, tout en équilibrant complexité et consommation d'énergie. Une application majeure de

cette absorption concerne la PLS, où un ARIS (Absorbent RIS) améliore la sécurité en contrôlant et absorbant sélectivement les réflexions des signaux, compliquant leur interception [28]. Cette fonctionnalité est essentielle pour maintenir le taux de secret dans les systèmes sans fil.

2.2.3 Principe de fonctionnement du RIS

2.2.3.1 Concept de base de fonctionnement

Le principe de fonctionnement du RIS repose sur sa capacité à manipuler les ondes EM incidentes pour optimiser la communication sans fil. Composé d'une surface dotée de nombreux éléments passifs ajustant dynamiquement la phase des signaux réfléchis, le RIS est programmé pour diriger les signaux vers un récepteur ou un point d'accès spécifique. En ajustant les déphasages, il maximise les interférences constructives, réduit les interférences non désirées et améliore la qualité du signal, le tout sans nécessiter de puissance supplémentaire pour l'amplification, ce qui le rend très économe en énergie. Placé dans des environnements complexes comme les zones urbaines ou intérieures, le RIS transforme le canal sans fil en une entité programmable, améliorant couverture, capacité et efficacité énergétique.

2.2.3.2 Contrôle du déphasage

Du point de vue de l'optique géométrique, le phénomène de réflexion anormale provoqué par les RIS sur la Fig.2.2 peut être décrit en utilisant la loi de réflexion généralisée dans (2.1).

$$\sin(\theta_r) - \sin(\theta_i) = \frac{2\pi\Delta x_{mn}}{\lambda} \Delta\Phi_{mn} \quad (2.1)$$

Où:

- θ_i : Angle d'incidence

- θ_r : Angle de réflexion
- λ : Longueur d'onde
- Δx_{mn} : Distance entre les cellules unitaires
- $\Delta \Phi_{mn}$: Différence de déphasage entre les UC (Unit Cell) voisines

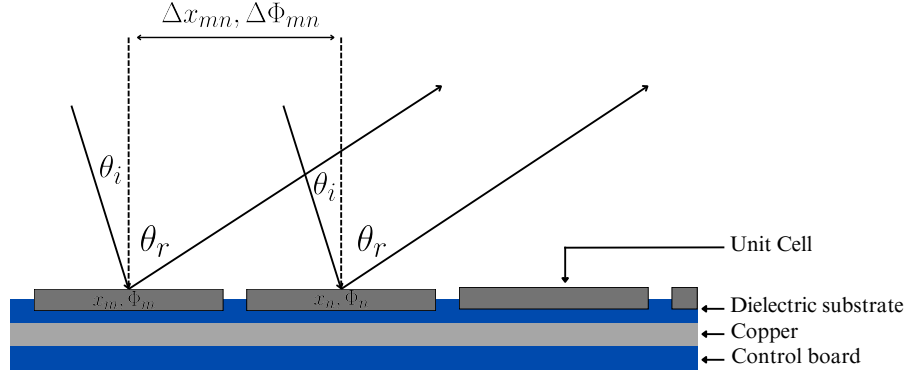


Figure 2.2 Illustration du mécanisme de réflexion du RIS et de la configuration des déphasages

Cette loi définit la différence de déphasage ($\Delta \Phi$) requise pour un contrôle précis de la direction et de la focalisation des ondes. En ajustant indépendamment le déphasage de chaque UC, l'onde EM réfléchie peut être orientée ou focalisée. La différence de déphasage entre les UC adjacents pour refléter une onde incidente à l'angle θ_r est celle indiquée dans (2.2).

$$\Delta \Phi_{mn} = \frac{2\pi \Delta x_{mn}}{\lambda} (\sin(\theta_r) - \sin(\theta_i)) \quad (2.2)$$

Le coefficient de réflexion, $\Gamma(\omega)$, est lié à l'impédance UC Z_{in} et est exprimé par (2.3).

$$\Gamma(\omega) = \frac{Z_{in} - Z_o}{Z_{in} + Z_o} \quad (2.3)$$

Où:

- Z_{in} est l'impédance d'entrée de l'UC.
- Z_o est l'impédance caractéristique de l'espace libre (377 ohms).

La différence de phase entre les UC voisines est alors définie dans (2.4).

$$\Delta\Phi_{mn} = \Phi_m(\Gamma) - \Phi_n(\Gamma) \quad (2.4)$$

En contrôlant l'impédance de chaque UC, le RIS peut diriger les faisceaux, focaliser les signaux ou appliquer une modulation par déphasage pour la transmission de données.

La plage de déphasage dans un RIS dépend en grande partie du type de mécanisme de contrôle utilisé dans les UC. Les déphasages discrets sont souvent obtenus grâce à l'utilisation de diodes PIN ou de commutateurs RF (Radio Frequency), qui fournissent un ensemble de valeurs fixes pour les déphasages. La quantification de phase sur un bit est couramment mise en œuvre à l'aide de ces composants, permettant deux états de phase distincts 0° et 180° . Cette approche simplifie la conception et réduit les coûts tout en permettant un contrôle efficace des ondes EM réfléchies pour des applications telles que la formation des faisceaux et la modulation des signaux.

La différence de déphasage quantifiée pour de tels systèmes peut être exprimée comme dans (2.5).

$$\Delta\Phi_{mn} = \begin{cases} \frac{\pi}{2}, & \text{pour un décalage demi-onde} \\ 0, & \text{Sinon} \end{cases} \quad (2.5)$$

À l'aide de diodes PIN et de commutateurs RF, les RIS permettent un contrôle précis du déphasage, crucial pour modifier le comportement du signal réfléchi et optimiser les scénarios de communication.

2.2.4 Les applications du RIS

La technologie RIS est devenue un élément clé dans diverses applications du monde réel en raison de sa capacité à révolutionner l'optimisation du signal et à améliorer la communication. Sa polyvalence lui permet d'être utilisé dans plusieurs domaines, offrant des solutions personnalisées pour répondre à des besoins spécifiques, comme le montre la Fig.2.3, conçue spécifiquement pour ce mémoire. En commençant par MEC (Mobile

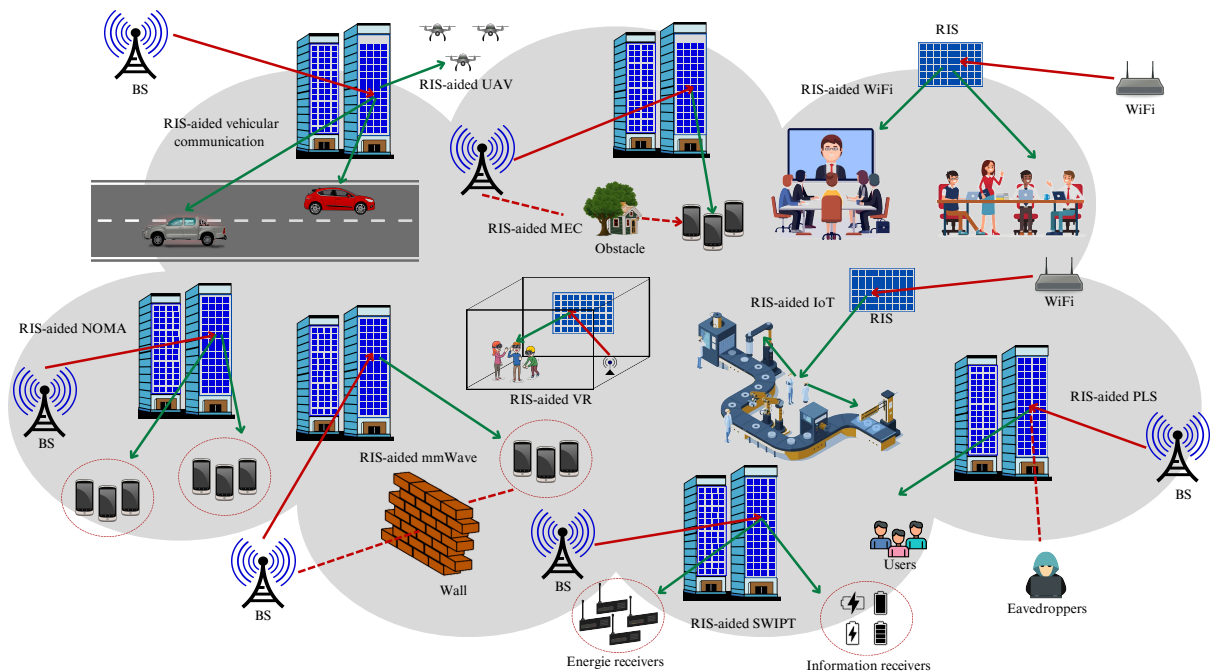


Figure 2.3 Les différents applications du RIS

Edge Computing), RIS améliore la fiabilité et l'efficacité grâce à une gestion avancée des ressources. Cependant, les appareils situés loin des nœuds MEC connaissent souvent de faibles taux de déchargement de données en raison de la perte de chemin [29]. Pour résoudre ce problème, un cadre MEC assisté par RIS a été introduit dans [30] et [31].

Dans les systèmes à mmWave, qui offrent des vitesses de données élevées et une faible latence cruciale pour les applications en temps réel, le RIS joue un rôle essentiel dans la

résolution des blocages de signaux causés par les problèmes de propagation. Les systèmes mmWave assistés par RIS, comme illustré dans [32], utilisent des réseaux de réflexion intelligents pour diriger efficacement les signaux et maintenir des connexions solides. Le RIS s'est également révélé utile pour les communications intérieures, où il peut être placé stratégiquement sur les murs ou les plafonds pour optimiser la couverture et la qualité du signal. Ceci est particulièrement avantageux pour les applications gourmandes en données telles que la réalité virtuelle, le streaming vidéo HD, les jeux et les téléconférences. Dans [33], des mesures pratiques ont démontré une couverture de signal améliorée pour les déploiements mono-utilisateur et une communication sécurisée pour les scénarios multi-utilisateurs.

Au-delà des systèmes de communication conventionnels, RIS s'est montré prometteur dans l'amélioration des réseaux IoT (Internet of Things). L'étude dans [34] examine comment RIS améliore les performances en cas d'évanouissement et d'ombrage, en fournissant des mesures de capacité, de taux de BER (Bit Error Rate) et de probabilité de panne. Il met en évidence des gains et des compromis significatifs entre la mise à l'échelle des cellules RIS et l'augmentation de la puissance de transmission dans des environnements difficiles.

Dans le contexte des réseaux de véhicules pour la conduite autonome, RIS a joué un rôle déterminant dans l'amélioration des communications V2X (Vehicle-to-Everything) en relevant des défis tels que le double évanouissement, le vieillissement des canaux et la grande mobilité. Cela conduit à des connexions plus fiables, à une formation de faisceau efficace et à une réduction de la surcharge de signalisation, comme le souligne [35]. De même, dans les systèmes de communication UAV (Unmanned Aerial Vehicle), RIS améliore la transmission du signal aux utilisateurs au sol en améliorant la couverture

et la fiabilité. L'étude en [36] analyse des indicateurs de performances clés tels que la probabilité de panne et la qualité du signal, tout en explorant également le positionnement et l'altitude optimaux des drones.

Une autre application importante des RIS consiste à renforcer la sécurité des systèmes de communication. Les techniques traditionnelles de PLS (Physical Layer Security) rencontrent souvent des difficultés lorsque les utilisateurs et les espions rencontrent des conditions de canal similaires. Dans [37], un système PLS assisté par RIS améliore les taux de secret en optimisant la formation de faisceaux de transmission et de réflexion, renforçant ainsi les signaux légitimes tout en supprimant les écoutes indiscretes. RIS a également été intégré aux systèmes NOMA (Non-Orthogonal Multiple Access), permettant à plusieurs utilisateurs de partager simultanément les mêmes ressources de fréquence. L'étude dans [38] démontre que le NOMA assisté par RIS surpasse le NOMA conventionnel, en particulier lorsque le RIS est placé de manière optimale.

Enfin, RIS améliore les systèmes SWIPT (Simultaneous Wireless Information and Power Transfer) en améliorant les performances MIMO (Multiple-Input Multiple-Output) multi-utilisateurs, permettant une transmission efficace des données et une récupération d'énergie avec une perte de signal minimale [39].

2.3 Travaux antérieurs

La suite examine les contributions des travaux précédents sur l'optimisation des RIS, en mettant en lumière les approches développées pour améliorer leurs performances et leur efficacité dans divers scénarios. Les techniques d'optimisation avancées sont essentielles pour gérer la complexité des RIS et exploiter pleinement leur potentiel. Motivé par l'importance de ces approches, ce travail explore diverses stratégies de conception,

d'analyse, d'application et d'optimisation dans le contexte des communications sans fil assistées par RIS, notamment à travers les algorithmes évolutionnaires et les méthodes de ML (Machine Learning). Les RIS offrent de riches perspectives de recherche, notamment en modélisation et estimation des canaux, traitement du signal, analyse de puissance, formation de faisceaux passifs et conception matérielle.

L'application de l'DRL (Deep Reinforcement Learning) aux réseaux sans fil assistés par RIS a entraîné des progrès significatifs dans l'optimisation des réseaux. Les méthodes basées sur le DRL excellent dans la résolution de problèmes tels que l'efficacité énergétique, la variabilité des canaux et l'adaptabilité en temps réel, surpassant souvent les approches traditionnelles. Par exemple, dans [40], DRL est utilisé pour améliorer l'efficacité énergétique en optimisant conjointement la puissance de la BS (Base Station) et les déphasages RIS, en s'adaptant dynamiquement aux changements environnementaux. De même, [41] applique le DRL pour minimiser la consommation d'énergie dans les systèmes MIMO en optimisant simultanément la formation de faisceaux et les configurations RIS. DRL améliore également les performances des réseaux assistés par drones, comme le montrent [42] et [43], où il optimise les déphasages RIS et les trajectoires des drones, réduisant ainsi la latence et améliorant l'efficacité énergétique. En complément de ces méthodes, les GA (Genetic Algorithm) se sont révélés robustes dans l'optimisation des systèmes assistés par RIS, en particulier pour les tâches d'optimisation plus complexes. Par exemple, [44] utilise GA pour optimiser les débits réalisables dans les systèmes RIS multi-paires, gérant efficacement les déphasages continus et discrets. De même, [45] introduit un NGA (Niching Genetic Algorithm) pour la formation de faisceaux conjointe et l'optimisation RIS, surmontant la complexité des configurations RIS à grande échelle. Ces articles démontrent la capacité de GA à naviguer dans des paysages d'optimisation complexes, ce qui en fait un outil précieux pour améliorer

les performances des réseaux assistés par RIS. S'appuyant sur les atouts de GA et de l'apprentissage automatique, cet article [46] explore l'utilisation d'algorithmes heuristiques pour optimiser les réseaux sans fil assistés par RIS, une technologie clé pour activer les SRE dans les réseaux 6G. La complexité de la gestion de ces réseaux présente des défis importants. Pour résoudre ce problème, les auteurs proposent l'utilisation d'algorithmes gloutons, d'algorithmes méta-heuristiques comme GA et PSO (Particle Swarm Optimization), et de la théorie de l'appariement pour fournir des solutions efficaces et peu complexes. Une contribution majeure de [46] est l'intégration de ces algorithmes heuristiques avec des techniques de ML, en particulier le DRL assisté par heuristique. Cette approche hybride accélère la convergence et permet d'atteindre des débits de données plus élevés, offrant ainsi une nouvelle perspective sur l'optimisation du contrôle RIS dans les environnements sans fil dynamiques.

En développant l'intégration de l'GA et du DL, [47] introduit un algorithme d'apprentissage G-MAQ (Genetic Multi-Agent Q-learning) conçu pour améliorer les systèmes MIMO en duplex intégral avec STAR-RIS double face. L'approche G-MAQ combine le Q-learning, où plusieurs agents apprennent à optimiser la formation de faisceaux et les configurations, avec GA, qui affine le processus d'apprentissage. Cette méthode hybride améliore considérablement les débits de données pour les liaisons montantes et descendantes, surpassant les techniques traditionnelles et augmentant l'efficacité globale du système dans les communications full-duplex.

2.4 Contribution

La recherche récente a démontré l'efficacité des approches basées sur le DRL et les GA pour optimiser les réseaux sans fil assistés par RIS, en adressant des problèmes tels que l'efficacité énergétique, l'adaptabilité en temps réel et la complexité des configurations

RIS. Toutefois, un potentiel reste à explorer dans la synergie entre les approches adversariales et évolutionnaires.

La contribution majeure de ce mémoire repose sur l'intégration novatrice des algorithmes de AB (Adversarial Bandit) avec les GA pour l'optimisation des RIS. Contrairement aux approches traditionnelles, cette méthode hybride associe la capacité des AB à explorer efficacement des environnements dynamiques et la puissance des GA pour affiner les configurations et garantir une optimisation robuste.

Les approches DRL, telles que celles de [40] et [41], sont performantes pour l'adaptabilité en temps réel mais peuvent nécessiter des périodes d'entraînement longues et être sensibles aux environnements dynamiques. De leur côté, les GA, comme dans [44] et [45], naviguent efficacement dans des paysages d'optimisation complexes mais tendent à converger plus lentement dans certains contextes, notamment lorsque l'espace de recherche est très vaste [13].

En intégrant le AB et le GA, notre méthode surmonte ces limitations. Les AB explorent rapidement les configurations prometteuses, tandis que les GA affinent ces configurations, accélérant ainsi la convergence et augmentant la robustesse dans des environnements variables. Cette combinaison réduit également les besoins computationnels par rapport aux approches DRL ou GA isolées. Cette approche hybride améliore les performances en termes de débit de données, d'efficacité énergétique et de gestion des interférences, tout en offrant une solution adaptée aux scénarios complexes, tels que les réseaux multi-utilisateurs ou les environnements densément obstrués. Elle représente une avancée significative pour l'optimisation des RIS dans les réseaux 6G, répondant aux défis des environnements radio intelligents de nouvelle génération.

2.5 Contexte institutionnel et continuité des travaux

Plusieurs projets menés au sein du Laboratoire des Signaux et Systèmes Intégrés (LSSI) portent sur l'intégration des RIS dans les systèmes de communication sans fil. Ces travaux permettent de situer le présent projet dans une continuité d'initiatives axées sur l'expérimentation concrète et l'optimisation des réseaux :

- **5G assisté par RIS (RIS-aided 5G)** : Intégration d'un RIS dans un environnement 5G [48] pour tester, analyser et démontrer ses impacts sur les performances du lien radio — recherche en cours menée par Sifeddine Salmi et Nada Belhadj Ltaief.
- **Déploiement de plateformes SDR** : Mise en œuvre de transmetteurs et récepteurs SDR pour la configuration et l'analyse de systèmes en conditions réelles [49] — recherche en cours menée par Mohammadkarim Shafieian.
- **Amélioration des performances Wi-Fi** : Allocation des ressources assistée par RIS [13] pour optimiser les performances dans les environnements Wi-Fi — recherche en cours menée par Mohammadkarim Shafieian et Nada Belhadj Ltaief.
- **Optimisation des réseaux 6G** : Intégration de RIS dans une architecture Open RAN à l'aide de FlexRIC pour améliorer les performances des réseaux 6G [50] — recherche en cours menée par Sifeddine Salmi et Nada Belhadj Ltaief.
- **RIS multifonction avec capacités de détection intégrée** : Considération d'un RIS multifonction combinant réflexion intelligente et détection pour étendre les fonctionnalités des réseaux sans fil [51] — recherche en cours menée par Mohammadkarim Shafieian.

2.6 Conclusion

Ce chapitre a exploré les concepts clés des RIS, notamment leur architecture, modes de fonctionnement et applications, soulignant leur rôle dans l'optimisation des réseaux sans

fil. Les travaux antérieurs ont démontré l'efficacité des approches basées sur le DRL et les GA, mais des défis subsistent en termes de robustesse et de convergence. La contribution de cette recherche propose une méthode hybride innovante, combinant les AB et GA, pour surmonter ces limitations. Cette approche promet des performances renforcées et une meilleure adaptabilité dans les réseaux 6G. Le chapitre établit ainsi un cadre pour les contributions et perspectives de ce projet.

Chapitre 3 - Méthodologie

3.1 Introduction

Ce chapitre présente le principe de contrôle du RIS et les principales stratégies d'optimisation pour la configuration du RIS dans un système de communication à utilisateur unique. L'objectif est de maximiser les performances du système en ajustant efficacement les configurations du RIS.

3.2 Cadre du système et formulation du problème

Dans ce cadre système, le RIS, placé stratégiquement entre l'émetteur et le récepteur, est composé d'éléments passifs configurables individuellement. Ces éléments permettent de contrôler la phase et l'amplitude des signaux réfléchis, améliorant ainsi la communication en contournant les obstacles, réduisant les interférences et concentrant l'énergie du signal vers le récepteur. Sans le RIS, le signal circulerait directement de l'émetteur au récepteur, ce qui pourrait entraîner une dégradation du signal en raison d'obstacles, d'interférences ou de distance.

Le signal reçu sans RIS peut être exprimé par (3.1).

$$y = h_{LoS}s + n \quad (3.1)$$

où :

- h_{LoS} est le gain de canal direct entre l'émetteur et le récepteur causé par la LoS.
- s est le signal transmis.
- n est le bruit au niveau du récepteur.

Avec l'intégration du RIS, le signal reçu y est une combinaison du signal direct et des

signaux réfléchis par les éléments RIS, modulés en phase et en amplitude. Cela peut être représenté par (3.2).

$$y = h_{LoS} \cdot s + \sum_{i=1}^N h_{r,i}^{\text{Tx-RIS}} \cdot \phi_i \cdot h_{r,i}^{\text{RIS-Rx}} \cdot s + n \quad (3.2)$$

où :

- $h_{r,i}^{\text{Tx-RIS}}$ est le gain de canal de l'émetteur vers le i -ième élément RIS.
- $\phi_i = \alpha_i e^{j\theta_i}$ représente le coefficient de réflexion du i -ième élément RIS, avec α_i comme amplitude et θ_i comme déphasage (valeurs possibles : 0° , 180°).
- $h_{r,i}^{\text{RIS-Rx}}$ est le gain de canal du i -ième élément RIS vers le récepteur.
- N est le nombre total d'éléments RIS.

Dans notre configuration, la communication directe LoS entre l'émetteur et le récepteur est négligée en raison de l'utilisation d'antennes cornet hautement directionnelles, orientées exclusivement vers le RIS. Par conséquent, le signal reçu repose uniquement sur les contributions des signaux réfléchis par les éléments RIS, et la composante directe est ignorée. De plus, le bruit est omis dans cette analyse en raison d'un SNR (Signal-to-Noise Ratio) élevé, où la puissance du signal dépasse largement celle du bruit, rendant son impact négligeable.

Le signal reçu peut alors être exprimé dans (3.3).

$$y = \sum_{i=1}^N h_{r,i}^{\text{Tx-RIS}} \cdot \phi_i \cdot h_{r,i}^{\text{RIS-Rx}} \cdot s \quad (3.3)$$

L'objectif principal est d'optimiser la configuration du RIS pour maximiser le RSSI (Received Signal Strength Indicator) au niveau du récepteur. Le RSSI, qui reflète la puissance du signal reçu, est crucial pour la qualité et la fiabilité de la communication.

Il est donné par (3.4).

$$RSSI = \left| \sum_{i=1}^N h_{r,i}^{\text{Tx-RIS}} \cdot \phi_i \cdot h_{r,i}^{\text{RIS-Rx}} \cdot s \right|^2 \quad (3.4)$$

Le problème d'optimisation peut être formulé dans (3.5).

$$\phi^* = \arg \max_{\phi} RSSI \quad (3.5)$$

où ϕ^* est la configuration optimale du RIS.

L'optimisation vise à ajuster les coefficients de réflexion ϕ_i afin que les signaux réfléchis par les éléments RIS soient combinés de manière constructive au niveau du récepteur, maximisant ainsi le RSSI. Cette approche exploite la programmabilité du RIS pour s'adapter dynamiquement à l'environnement, transformant efficacement les conditions de propagation afin d'améliorer la réception du signal.

3.3 Ressources matérielles et logiciels pour le fonctionnement du système

Cette section présente les différents éléments matériels et logiciels essentiels au fonctionnement du système. Chaque composant joue un rôle clé dans la mise en œuvre et l'optimisation des performances. Les ressources matérielles incluent les dispositifs de contrôle, de communication et de traitement des signaux, tandis que les outils logiciels assurent la gestion des interactions et le transfert des données. Ensemble, ces éléments constituent l'infrastructure nécessaire à l'étude et à l'évaluation du système.

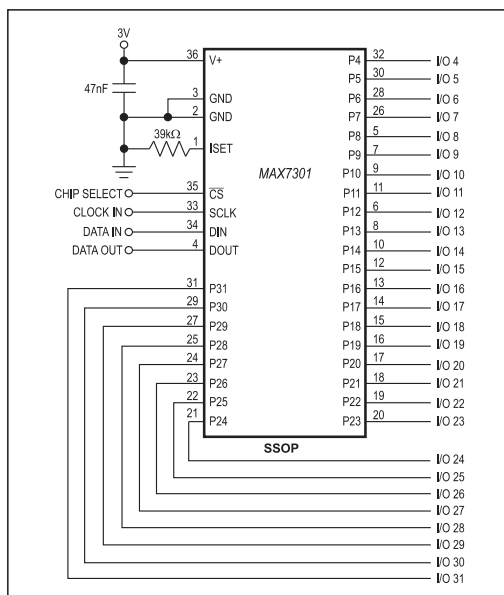


Figure 3.1 Les pins du MAX7301

3.3.1 Éléments constitutifs du système

3.3.1.1 MAX7301: Extenseur GPIO

Le MAX7301 est un extenseur GPIO (General Purpose Input Output) compact et polyvalent, conçu pour étendre les capacités d'E/S des microprocesseurs via une interface SPI à quatre fils. Il offre jusqu'à 28 ports configurables individuellement représentés dans la figure 3.1, pouvant fonctionner comme sorties logiques push-pull (4,5 mA en sortie et 10 mA en entrée) ou comme entrées logiques à déclenchement Schmitt avec résistances de pull-up internes.

Cette flexibilité rend le MAX7301 adapté à diverses applications, comme le contrôle d'antennes ou de LED (Light-Emitting Diode). Il intègre également une fonction de détection de transitions sur sept ports, permettant une gestion d'événements efficace via

des interruptions. Conçu pour une faible consommation d'énergie, il dispose d'un mode veille qui réduit le courant à 11 μA .

En outre, le MAX7301 peut être chaîné en série pour connecter plusieurs dispositifs. Cela permet d'augmenter significativement le nombre de GPIO contrôlables tout en simplifiant la communication via le protocole SPI. Ces caractéristiques en font un composant idéal pour les systèmes embarqués nécessitant une extension d'E/S évolutive et à faible consommation, comme l'implémentation du RIS.

3.3.1.2 *Module Software-Defined Radio*

Une radio définie par logiciel ou SDR (Software-Defined Radio) est une technologie qui remplace les composants matériels traditionnels d'un système radio par des logiciels capables de traiter des signaux numériques. Cela permet une flexibilité et une reconfigurabilité améliorée, rendant possible l'émulation de divers systèmes de communication sur un seul appareil.

L'ADALM-PLUTO, également connu sous le nom de PlutoSDR, est un exemple accessible et performant de SDR conçu par Analog Devices. Ce dispositif éducatif compact est idéal pour explorer les concepts de communication sans fil et de traitement des signaux. Il couvre une plage de fréquences de 325 MHz à 3,8 GHz, avec une bande passante instantanée jusqu'à 20 MHz. Intégrant un émetteur et un récepteur en duplex intégral, le PlutoSDR permet l'expérimentation sur des systèmes de transmission et réception. Alimenté par le transcepteur RF AD9363 et le FPGA Xilinx Zynq Z-7010, il est compatible avec des outils comme MATLAB et Simulink, offrant une plateforme flexible pour les projets éducatifs et de prototypage.

Dans ce projet, deux modules PlutoSDR sont utilisés, l'un en émission pour transmettre

un signal modulé vers la surface RIS, et l'autre en réception pour mesurer la puissance du signal réfléchi, constituant ainsi la boucle de rétroaction essentielle à l'optimisation.

3.3.1.3 *Antenne unidirectionnel*

Une antenne cornet est un type d'antenne unidirectionnelle qui utilise une structure en forme de cornet pour diriger les ondes EM. Cette conception permet une adaptation progressive entre le guide d'ondes et l'espace libre, améliorant ainsi l'efficacité du rayonnement et la directivité. Les antennes cornets sont couramment utilisées dans les applications hyperfréquences, notamment pour les mesures en chambre anéchoïque et les systèmes radar, en raison de leur large bande passante et de leur gain élevé. Dans ce projet, Les antennes cornets sont utilisées pour focaliser le signal en direction du RIS à l'émission, puis pour capter efficacement le signal réfléchi à la réception, assurant une liaison sans LoS.

3.3.1.4 *Microcontrôleur*

La carte de développement STM32 Nucleo-144 avec le microcontrôleur STM32F746ZG est une plateforme puissante pour les applications embarquées. Basée sur un cœur ARM Cortex-M7, elle offre des performances élevées avec 1 Mo de mémoire Flash, 320 Ko de RAM et une fréquence jusqu'à 216 MHz. Compatible avec les connecteurs Arduino™ Uno V3, ST Zio et ST morpho, elle prend en charge des interfaces comme UART, SPI, I2C et USB. Elle intègre également un débogueur ST-LINK/V2-1, facilitant le développement. Idéale pour des projets IoT, d'automatisation ou d'applications industrielles. Dans l'architecture du système, le microcontrôleur STM32 agit comme l'unité de commande centrale du RIS, appliquant directement les configurations nécessaires pour piloter les éléments réfléchissants.

3.3.2 Protocoles utilisés pour l'interconnexion

3.3.2.1 TCP/IP

Le protocole TCP/IP (Transmission Control Protocol/Internet Protocol) est un ensemble de règles permettant la communication entre appareils sur un réseau, y compris Internet. TCP gère la fiabilité des transmissions en assurant que les données sont reçues dans le bon ordre et sans perte, tandis qu'IP est responsable de l'adressage et de l'acheminement des paquets vers leur destination. Ce protocole, organisé en couches, offre une grande flexibilité et s'adapte à différents types de réseaux, garantissant des échanges rapides et fiables.

3.3.2.2 MQTT

Le protocole MQTT (Message Queuing Telemetry Transport), est un protocole de communication M2M (Machine-to-Machine) basé sur TCP/IP, conçu pour transmettre des messages de manière légère et efficace. Il repose sur un modèle publication/abonnement où les clients se connectent à un serveur central appelé "broker". Les messages publiés sont organisés par sujets hiérarchiques, permettant aux abonnés de recevoir uniquement les informations pertinentes. Une connexion TCP/IP est établie entre le client et le broker via les ports 1883 ou 8883 (avec TLS), et reste active jusqu'à sa fermeture par le client. Les échanges utilisent des commandes spécifiques : PUBLISH pour envoyer des messages, SUBSCRIBE et UNSUBSCRIBE pour gérer les abonnements, et PINGREQ/PINGRESP pour vérifier l'état de la connexion. MQTT est ainsi une solution fiable, sécurisée et adaptée aux environnements IoT nécessitant une communication efficace et peu gourmande en ressources.

3.3.2.3 SPI

SPI est une interface synchrone, full-duplex, largement utilisée pour la communication entre microcontrôleurs et circuits intégrés. Basée sur une configuration maître-esclave, elle permet la transmission simultanée de données entre le maître et les esclaves, synchronisée avec le front montant ou descendant d'une horloge. L'interface SPI utilise quatre fils principaux : Clock (SCLK), Chip Select (CS), Master-Out Slave-In (MOSI) et Master-In Slave-Out (MISO), comme illustré dans la Figure 3.2.

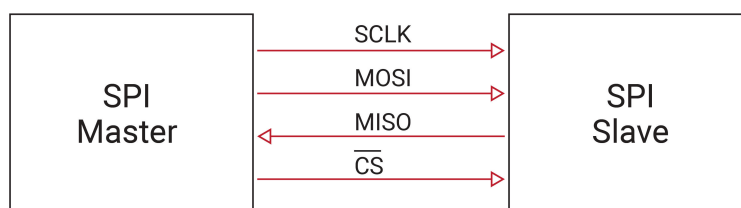


Figure 3.2 Configuration SPI avec maître et esclave

Dans une configuration SPI, le maître génère le signal SCLK et contrôle la communication avec les esclaves via le signal CS, généralement actif à l'état bas. Lorsqu'il y a plusieurs esclaves, chaque esclave nécessite un signal CS distinct pour être adressé. Les données sont transmises via MOSI (du maître vers l'esclave) et MISO (de l'esclave vers le maître). Ce mécanisme garantit une communication efficace et synchronisée, essentielle pour les applications embarquées complexes.

3.3.3 Langages de Programmation et Outils Logiciels

Ce projet utilise trois outils clés, le C embarqué pour un contrôle matériel précis, Python pour le prototypage rapide et l'automatisation, et GNU Radio pour la configuration et l'analyse des transmissions SDR. Ensemble, ils assurent la programmation embarquée, la simulation et l'intégration du RIS dans divers scénarios de communication.

3.4 Conception et contrôle RIS

Le RIS utilisé dans ce projet est un dispositif personnalisé. Il permet de configurer dynamiquement les éléments pour optimiser les performances. Cette conception vise à répondre aux exigences des scénarios de communication avancés.

3.4.1 RIS personnalisé pour la réalisation du projet

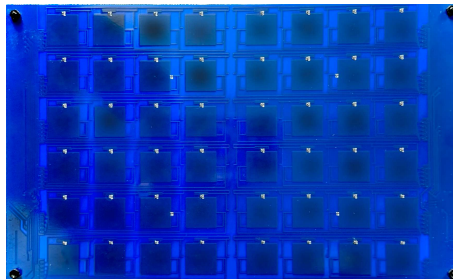


Figure 3.3 RIS personnalisé

Le RIS développé présente une structure soigneusement conçue visant à améliorer son application pratique dans les systèmes de communication sans fil. La figure 3.3 illustre la disposition physique du RIS réel, qui comprend 48 UC.

Le schéma électronique du RIS est représenté dans la figure 3.4, les dimensions totale du RIS sont de 29,5 cm x 18,2 cm. Les deux panneaux latéraux du RIS, mesurant chacun 18,2 cm x 2,75 cm, contiennent les circuits de commande qui ajustent les déphasages des 24 unités cellulaires de chaque côté. Les 24 UC de chaque côté sont espacées de 20 mm. Chaque UC mesure 20 mm x 20 mm et est positionné à égale distance avec un espacement de 10 mm de tous les côtés adjacents. Cette disposition précise garantit une réflexion et une adaptabilité optimales du signal, soulignant la capacité du RIS à répondre aux exigences des systèmes de communication sans fil modernes.

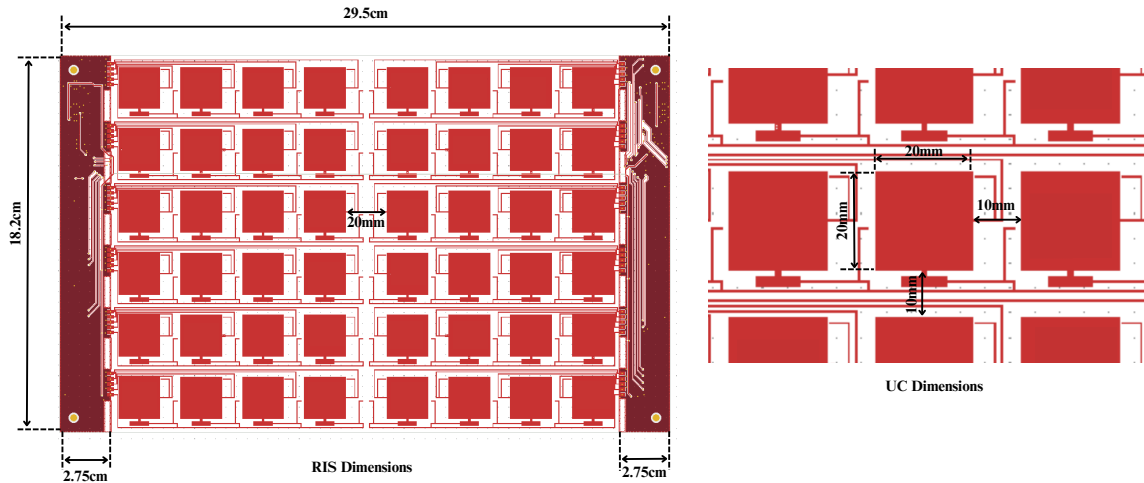


Figure 3.4 Les dimensions du RIS et des UC

3.4.2 Couches matérielles du RIS

Pour mieux comprendre la structure, la Fig.3.5 présente une décomposition détaillée de l'architecture à trois couches. La première et la plus importante couche est le réseau d'antennes, composé de 48 cellules soigneusement disposées selon une géométrie périodique. Ces cellules, associées à des diodes PIN intégrées, constituent l'épine dorsale de la capacité du RIS à manipuler les ondes EM. Chaque unité d'antenne peut ajuster la phase des ondes entrantes en modifiant l'état de la diode PIN associée, offrant ainsi un contrôle précis sur la réflexion et l'orientation des signaux. Ces diodes jouent un rôle fondamental dans le contrôle des propriétés de réflexion du RIS, alternant entre différents états opérationnels grâce au contrôle de la tension de polarisation. La deuxième couche, constituée de cuivre, agit comme le chemin conducteur principal du système, connectant les différents composants avec précision. Cette couche garantit un flux de signaux électriques avec des pertes minimales, maximisant l'efficacité du traitement des signaux par le RIS. La troisième et dernière couche regroupe divers composants passifs, notamment des résistances, des condensateurs, etc., qui fonctionnent conjointement avec

les diodes PIN pour ajuster finement les caractéristiques électriques du RIS. De plus, des indicateurs LED sont associés à chaque UC pour fournir un retour d'information en temps réel sur l'état opérationnel, signalant si une UC est active ou inactive. Cela apporte une fonctionnalité clé au système de contrôle, facilitant une configuration et une surveillance en temps réel du RIS pour s'adapter aux évolutions des conditions sans fil.

3.4.3 Contrôle de déphasage basé sur une diode PIN

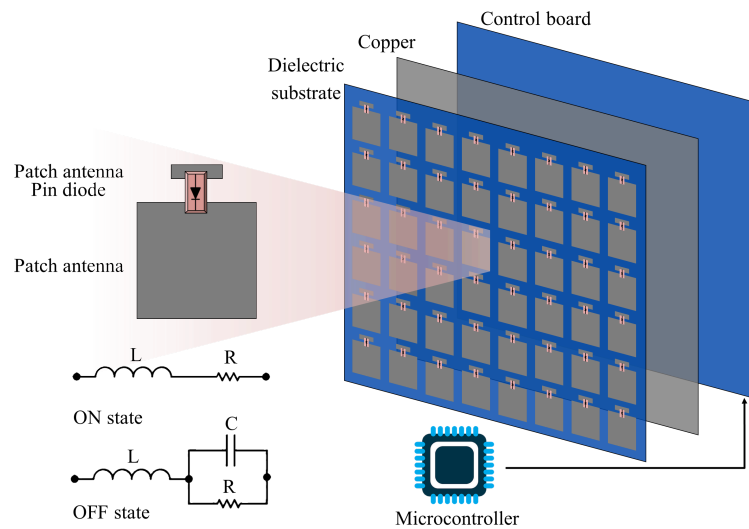


Figure 3.5 Architecture matérielle du RIS

Les diodes PIN sont essentielles pour commuter les éléments du RIS entre les états ON et OFF, modifiant ainsi l'impédance de chaque UC. À l'état ON, le circuit équivalent se comporte comme un circuit série RL avec un élément à faible résistance, offrant une impédance minimale et permettant aux ondes EM de se réfléchir avec un déphasage minimal. À l'inverse, à l'état OFF, le circuit équivalent passe à une configuration parallèle RLC, introduisant une impédance plus élevée et un déphasage plus important. Ces transitions sont contrôlées en faisant varier la tension de polarisation fournie aux diodes, comme illustré sur la Fig.3.5. Ce mécanisme de commutation précis offre la flexibilité nécessaire pour orienter, focaliser ou manipuler le signal réfléchi d'une manière qui était

auparavant impossible à réaliser en utilisant des surfaces statiques.

En utilisant une modélisation de circuit équivalente, nous pouvons prédire et ajuster avec précision le comportement du RIS dans différentes conditions de fonctionnement. Cette adaptabilité est essentielle pour les applications sans fil traitées en temps réel dans le cadre de ce projet, c'est-à-dire à partir de données acquises, analysées et optimisées dynamiquement pendant l'expérimentation, sans recourir à une base de données préexistante. Cette capacité de réponse instantanée permet d'adapter continuellement les configurations du RIS à l'évolution des conditions environnementales, afin de maintenir une force et une directivité optimales du signal.

Cette architecture matérielle, combiné avec la configuration contrôlée par logiciel du RIS, constitue une solution avancée pour les réseaux de communication de nouvelle génération, y compris la 5G et au-delà. Grâce à sa flexibilité, sa précision et son efficacité, cette conception du RIS offre un potentiel considérable pour améliorer la couverture sans fil, optimiser la qualité du signal et réduire les interférences dans divers environnements.

3.4.4 Contrôle et configuration du RIS

Le circuit de contrôle intégré dans les panneaux latéraux du RIS joue un rôle crucial dans son fonctionnement. Il comprend des résistances pour réguler les courants, des condensateurs destinés à stabiliser les tensions, et des diodes. La configuration et le contrôle du RIS reposent sur une architecture matérielle composée de composants clés tels que des extenseurs GPIO et un protocole de communication efficace. Ce système assure la gestion individuelle des éléments du RIS, garantissant une configuration précise et adaptée aux besoins des expérimentations. Les éléments matériels et protocolaires impliqués dans ce processus sont décrits ci-dessous.

3.4.5 Le contrôle du RIS via MAX7301

La conception du PCB du RIS repose sur un élément de déphasage à 1 bit. Le RIS comprend 48 éléments, chacun contrôlé individuellement pour appliquer un déphasage de 0° ou 180° en ajustant la tension de polarisation à 0V ou 3,3V. Pour permettre ce contrôle indépendant, le réseau nécessite un minimum de 48 GPIO. L'utilisation d'un module d'extension E/S tel que le MAX7301, basé sur le protocole SPI, est essentielle pour gérer efficacement les tensions de polarisation et simplifier le câblage.

Dans ce système SPI, le microcontrôleur agit comme maître, tandis que deux MAX7301 sont utilisés comme esclaves, chacun contrôlant 24 éléments du RIS. Chaque MAX7301 dispose de 28 ports E/S configurables, parmi lesquels certains sont utilisés pour le contrôle direct des tensions de polarisation. Un signal CS actif-bas dédié est utilisé pour adresser un esclave spécifique. Cependant, l'augmentation du nombre d'esclaves dans le réseau SPI entraîne une multiplication des lignes CS, rendant la disposition des circuits complexe. Pour pallier cette limitation, le chaînage en série offre une solution efficace.

Dans un système de chaînage en série, illustré en Annexe A, tous les esclaves reçoivent le même signal d'horloge SCLK et un seul signal CS actif-bas contrôle l'ensemble des esclaves. L'esclave 1 est le seul à recevoir directement les données de commande du microcontrôleur via le signal MOSI. Ces données sont enregistrées dans son registre à décalage interne. Tant que le signal CS reste à l'état bas, les données propagées depuis l'esclave 1 sont transmises via sa sortie DOUT à l'entrée DIN de l'esclave 2. De cette manière, l'esclave 2 enregistre les données transmises par l'esclave 1 dans son propre registre à décalage. Ce processus se poursuit à travers toute la chaîne jusqu'à ce que chaque esclave ait reçu sa commande.

Pour que le chaînage en série fonctionne correctement, chaque esclave doit respecter des exigences précises :

- Chaque commande doit être enregistrée via DIN pendant un cycle de commande défini par le nombre d'impulsions d'horloge nécessaires pour enregistrer une commande et propagée à DOUT pendant le cycle suivant, avec un délai d'un cycle de commande entre DIN et DOUT.
- Les esclaves doivent exécuter les commandes uniquement sur le front montant du signal CS actif-bas. Cela garantit que les données circulent à travers la chaîne tant que CS reste bas, mais ne sont exécutées qu'une fois que CS passe à l'état haut.

Ce mécanisme permet de contrôler simultanément plusieurs esclaves avec seulement trois signaux principaux : MOSI, SCLK et CS actif-bas. En outre, chaque esclave peut exécuter une commande différente en fonction des données qui lui sont spécifiques. Ce mode de communication optimise l'utilisation des lignes de signal, simplifie le câblage et assure un contrôle efficace des esclaves.

Enfin, les registres internes des MAX7301 jouent un rôle crucial dans ce processus. Chaque port E/S peut être configuré individuellement via les registres de configuration. Les données de commande sont ensuite propagées via le registre à décalage interne des esclaves, permettant une transmission fluide et ordonnée des commandes à travers la chaîne.

3.5 Principe du fonctionnement du système

Le fonctionnement du système repose sur une interaction fluide entre plusieurs composants, permettant de configurer et d'évaluer les performances des RIS. Chaque étape implique une coordination précise entre la génération des configurations, leur transmission au RIS, et la mesure des performances. Ce processus s'appuie sur des

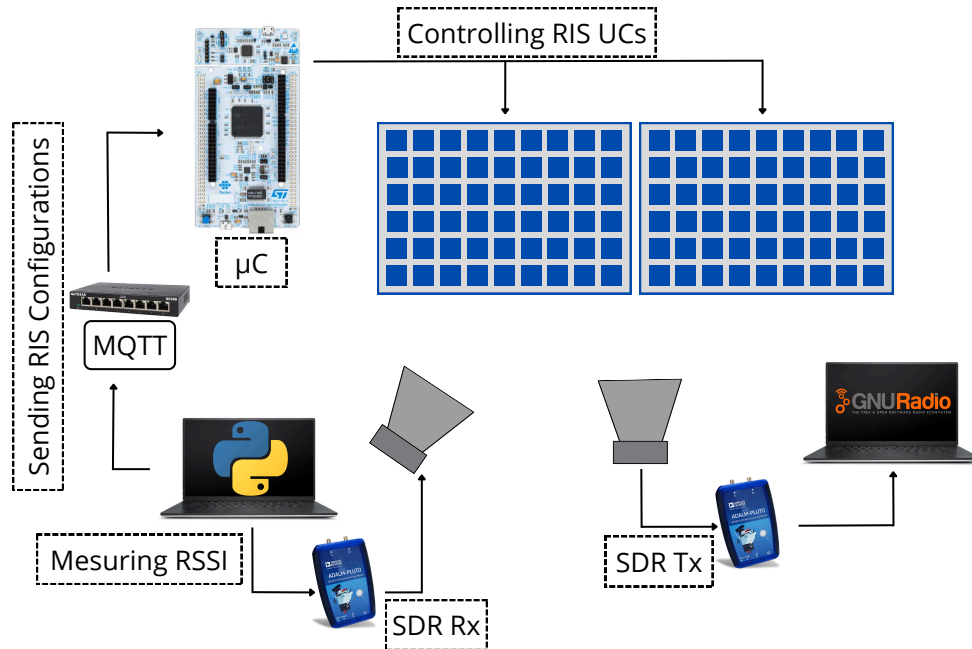


Figure 3.6 Description du fonctionnement de l'installation expérimentale

protocoles de communication adaptés, des interfaces matérielles spécifiques, et un flux de données contrôlé pour garantir une optimisation efficace. Le principe de fonctionnement du système est illustré dans la figure 3.6. Les configurations du RIS sont d'abord générées par un script Python qui les transmet au microcontrôleur exécutant du code C embarqué via le protocole MQTT utilisant une communication TCP/IP. Une fois reçues, ces configurations sont envoyées au RIS par le microcontrôleur via une interface SPI, comme décrit précédemment. Après l'application de chaque configuration au RIS, le RSSI est mesuré au niveau du récepteur pour évaluer les performances. Ce processus est répété pour explorer différentes configurations. Le transmetteur, programmé indépendamment sur un autre PC en utilisant GNU Radio, reste dédié à l'envoi continu des signaux vers le RIS pendant toute l'expérimentation. Ce flux coordonné garantit une optimisation efficace des performances du système.

3.6 Stratégies d'Optimisation pour le RIS

L'objectif de l'optimisation est de trouver la configuration RIS optimale pour améliorer la communication entre un émetteur et un récepteur. Cette section met en avant le AB assisté par GA, qui combine les avantages des algorithmes évolutifs et des stratégies d'exploration-exploitation. Le problème traité est de nature discrète, puisque chaque élément du RIS ne peut adopter que deux phases possibles 0° ou 180° . Cette contrainte binaire rend inapplicables les méthodes d'optimisation continue basées sur le calcul de gradients, et justifie l'utilisation d'algorithmes comme le GA ou le bandit adversarial, capables d'explorer efficacement un espace de solutions combinatoires. Les performances sont comparées avec le DRL, le DRL assisté par GA et le AB autonome. Chaque approche est décrite pour évaluer leur efficacité et leur pertinence.

3.6.1 GA comme approche de base

Les GA, appartenant aux méta-heuristiques évolutives, sont des méthodes itératives exploitant des stratégies heuristiques pour explorer efficacement de vastes espaces de recherche tout en restant peu complexes à concevoir [46]. Ces algorithmes utilisent des stratégies évolutives pour sélectionner des solutions d'élite et génèrent de nouvelles solutions via croisement et mutation, garantissant une amélioration progressive jusqu'à la convergence.

Contrairement à une recherche aléatoire, les GA exploitent les informations historiques pour orienter la recherche vers des solutions prometteuses, les rendant efficaces dans l'exploration d'espaces complexes [52]. De plus, ils offrent une alternative robuste aux techniques gourmandes, souvent limitées par les optima locaux. Leur performance dépend néanmoins de paramètres comme la taille de la population, le taux de mutation et le taux

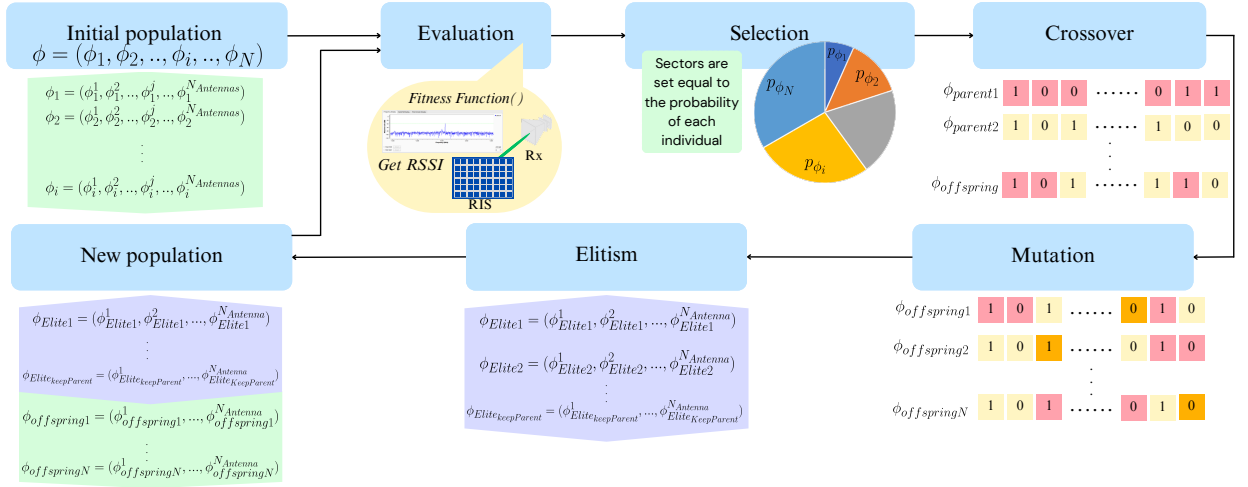


Figure 3.7 Architecture du RIS assistée par GA

de croisement, dont l'optimisation est cruciale pour des solutions quasi optimales.

Étant donné la relation non linéaire et complexe entre la configuration RIS et la force du signal reçu, GA est utilisé pour explorer l'espace de configuration RIS. Cet algorithme affine de manière itérative les coefficients de réflexion RIS pour maximiser le RSSI. La population du GA est composée de N individus, chacun représentant une configuration candidate pour le RIS. Ces configurations sont notées $\phi_1, \phi_2, \dots, \phi_i, \dots, \phi_N$, où chaque ϕ_i désigne un individu. Le processus d'initialisation, d'évaluation, de sélection, de croisement et de mutation des individus est détaillé dans l'algorithme 1.

1) Population initiale: une population initiale de $N_{\text{InitialPopulation}}$ individus notés $\phi_1, \phi_2, \dots, \phi_i, \dots, \phi_{N_{\text{InitialPopulation}}}$ est généré. Un individu ϕ_i est lui-même un vecteur binaire de taille N_{Antenna} , défini comme $\phi_i = (\phi_i^1, \phi_i^2, \dots, \phi_i^j, \dots, \phi_i^{N_{\text{Antenna}}})$, dans lequel chaque composante ϕ_i^j représente l'état actif ou inactif de l'élément j de la surface RIS. Chaque gène ϕ_i^j reçoit aléatoirement une valeur binaire, 0 ou 1, lors de l'initialisation. Ceci constitue l'ensemble initial de solutions candidates.

2) Évaluation: Chaque individu ϕ_i est évalué à l'aide de la fonction de fitness $f(\phi_i)$ qui

Algorithm 1 Algorithme génétique pour l'optimisation de la configuration du RIS

- 1: **Entrée** : Taille de la population initiale $N_{InitialPopulation}$, nombre de générations $N_{Generations}$, nombre de parents sélectionnés pour le croisement $N_{Parents}$, nombre de solutions par génération $N_{Solutions}$, nombre d'élites conservées N_{Elites} , taux de croisement, taux de mutation, espace des gènes, etc.
 - 2: **Initialisation** : Générer une population initiale de configurations RIS de taille $N_{InitialPopulation}$
 - 3: **for** $i = 1$ à $N_{Generations}$ **do**
 - 4: Évaluer l'aptitude (RSSI) de chaque individu dans la population
 - 5: Sélectionner $N_{Parents}$ parents en fonction de leur aptitude
 - 6: Appliquer le croisement pour créer des descendants jusqu'à atteindre $N_{Solutions} - N_{Elites}$
 - 7: Appliquer des mutations aux descendants
 - 8: Appliquer l'élitisme en ajoutant les N_{Elites} meilleurs individus aux descendants pour atteindre la taille requise $N_{Solutions}$
 - 9: Remplacer la population par la nouvelle génération
 - 10: **end for**
 - 11: **Sortie** : Meilleur RSSI et meilleure configuration RIS trouvée
-

Algorithm 2 Croisement uniforme pour l'algorithme génétique

- 1: **Entrée** : Individus parents $\phi_{parent1}, \phi_{parent2}$, Nombre d'individus nécessaires pour le croisement $N_t = N_{Solutions} - N_{Elites}$
 - 2: **Sortie** : Nouvelle génération de N_t individus descendants
 - 3: Initialiser le nombre de descendants générés : $n = 0$
 - 4: **while** $n < N_t$ **do** ▷ Continuer jusqu'à ce que le nombre de solutions par génération soit atteint
 - 5: **Pour chaque position de gène i dans le chromosome :**
 - 6: Générer un nombre aléatoire r entre 0 et 1
 - 7: **if** $r < 0.5$ **then**
 - 8: Définir $\phi_{offspring}^i = \phi_{parent1}^i$ ▷ Le gène est hérité du Parent 1
 - 9: **else**
 - 10: Définir $\phi_{offspring}^i = \phi_{parent2}^i$ ▷ Le gène est hérité du Parent 2
 - 11: **end if**
 - 12: Incrémenter n de 1 après la génération de chaque descendant
 - 13: **end while**
 - 14: **Retour** : Nouvelle génération de N_t individus descendants
-

mesure le RSSI.

3) Sélection: Au cours de chaque génération, les individus $\phi_1, \phi_2, \dots, \phi_{N_{\text{Solutions}}}$ de la population sont sélectionnés proportionnellement à leur fitness f_{ϕ_i} , où les solutions les plus adaptées ont de meilleures chances d'être choisies. Cette sélection probabiliste, donnée par l'équation (3.6), garantit la diversité au sein de la population en permettant même aux individus moins performants d'être sélectionnés, réduisant ainsi le risque de convergence prématurée vers des optima locaux.

$$p_{\phi_i} = \frac{f_{\phi_i}}{\sum_{j=1}^{N_{\text{Solutions}}} f_{\phi_j}} \quad (3.6)$$

où f_{ϕ_i} est l'aptitude de l'individu ϕ_i et $N_{\text{Solutions}}$ la taille de la population.

Dans notre cas, l'objectif est de maximiser le RSSI, bien que celui-ci prenne des valeurs négatives. Pour contourner cette contrainte et garantir que les probabilités de sélection restent positives et comparables, nous utilisons l'inverse $1/f_{\phi_j}$ comme mesure d'aptitude. Ce choix assure que les configurations générant des RSSI plus proches de zéro donc plus élevés sont favorisées. L'utilisation de $-f_{\phi_j}$ aurait produit des valeurs positives mais sans échelle relative exploitable, tandis que $f_{\phi_j}^2$ aurait accentué les écarts au détriment de la stabilité. Le recours à $1/f_{\phi_j}$ permet donc de préserver une relation monotone inversée avec le RSSI, tout en assurant la compatibilité avec une sélection probabiliste.

4) Croisement: À chaque génération, l'algorithme génère des descendants par un processus de croisement uniforme qui fusionne les matériaux génétiques obtenus des individus parents sélectionnés. Ce processus se poursuit jusqu'à ce que le nombre requis de solutions soit généré dans une génération donnée. Chaque descendant a une probabilité égale d'hériter d'un gène de l'un ou l'autre parent, favorisant ainsi la diversité et l'exploration de l'espace de recherche. Les étapes détaillées de la procédure de croisement

sont données dans l'algorithme 2.

5) Mutation: La fonction de mutation agit sur $(N_{\text{Solutions}} - N_{\text{Elites}}) \times L$ gènes, en excluant les individus élités. Ici, L représente le nombre de gènes par individu. Chaque gène a la possibilité de muter, c'est-à-dire de changer vers une valeur aléatoire dans l'intervalle $\{0, 1\}$ en fonction du taux de mutation. Cela permet d'explorer davantage l'espace de recherche et d'empêcher l'algorithme de rester bloqué dans un optimum local.

6) Élitisme: L'élitisme est une stratégie utilisée avec les GA pour conserver les meilleures solutions et éviter leur perte lors de la génération de nouveaux descendants. Sans élitisme, les meilleures solutions peuvent être perdues à cause de mutations aléatoires ou de croisements. Ce mécanisme retient un certain nombre de configurations élités, notées $\phi_{\text{Elite}_1}, \phi_{\text{Elite}_2}, \dots, \phi_{\text{Elite}_{\text{keepParent}}}$, où *keepParent* représente le nombre d'individus élités transmis à la génération suivante. Si les meilleurs descendants ne surpassent pas ces configurations élités, celles-ci sont conservées, assurant que les solutions performantes contribuent à la convergence de l'algorithme. Ce paramètre peut être ajusté pour équilibrer l'exploitation des solutions existantes et l'exploration de nouvelles configurations.

Le processus itératif du GA inclut les étapes de sélection, croisement, mutation et élitisme, jusqu'à ce qu'un critère de terminaison soit satisfait. Ces critères, comme un nombre maximal de générations ou des ratios spécifiques de croisement et mutation, garantissent un équilibre entre l'exploration de l'espace de recherche et l'efficacité du calcul, assurant que l'algorithme s'arrête lorsque d'autres itérations n'apporteraient plus d'améliorations significatives.

3.6.2 DRL et DRL assistés par GA comme méthodes comparatives

3.6.2.1 Préliminaires du DRL

Le RL est une méthode d'apprentissage où un agent apprend à prendre des décisions en interagissant avec un environnement et en recevant des retours sous forme de récompenses ou de pénalités. L'objectif est de développer une politique qui maximise la récompense cumulée sur le long terme, en identifiant les meilleures actions à entreprendre [53]. Le DRL étend le RL en utilisant des DNN (Deep Neural Network) pour traiter des espaces d'états et d'actions complexes. Ces réseaux permettent d'approximer soit la politique, soit la fonction de valeur, rendant les décisions de l'agent plus sophistiquées.

L'approche DQN (Deep Q-Network), une des plus populaires en DRL, combine le Q-learning avec des DNN pour estimer la valeur Q des actions. Cela permet à l'agent d'améliorer ses décisions en apprenant à maximiser les récompenses dans divers états. Le DQN repose sur des composants clés travaillant ensemble pour l'apprentissage par renforcement dans des environnements dynamiques :

Environnement: L'environnement est le système avec lequel l'agent interagit, générant observations et retours qui favorisent l'apprentissage, tout en évoluant selon les actions de l'agent.

Agent: l'agent interagit avec l'environnement afin d'atteindre un objectif. L'agent observe l'état actuel à chaque pas de temps et utilise les valeurs Q prédites émises par le DQN pour déterminer une action.

Etat: L'état s_t décrit la situation de l'environnement à l'étape t . Il fournit à l'agent les informations nécessaires pour décider quelle action entreprendre.

Action: L'action a_t est le choix de l'agent au moment t , déterminant comment l'agent influence l'environnement. L'ensemble de toutes les actions possibles est noté \mathcal{A} , et l'agent sélectionne une action qui maximise la valeur Q à chaque pas de temps.

Politique: La politique π guide l'agent dans le choix de ses actions. Dans le DQN, une politique greedy epsilon π_ϵ équilibre exploration et exploitation : l'agent choisit aléatoirement avec une probabilité ϵ au début, puis diminue ϵ au fil de l'apprentissage, privilégiant progressivement les actions avec les valeurs Q les plus élevées.

Récompense: La récompense r_t informe l'agent de l'efficacité de son action au moment t , renforçant les actions positives et décourageant les actions négatives pour l'aider à apprendre une politique optimale.

Fonction valeur: La fonction de valeur $Q(s_t, a_t)$ représente la récompense cumulée attendue pour l'action a_t dans l'état s_t au temps t , définie par (3.7).

$$Q(s_t, a_t) = \mathbb{E}_\pi [R_t \mid s_t = s, a_t = a] \quad (3.7)$$

où R_t , la récompense cumulée, est donnée par (3.8).

$$R_t = \sum_{\Delta t=0}^{T-t-1} \gamma^{\Delta t} r_{t+\Delta t+1} \quad (3.8)$$

avec :

- $\gamma \in [0, 1]$, le facteur d'actualisation.
- $r_{t+\Delta t+1}$, la récompense au pas de temps $t + \Delta t + 1$.

La fonction Q évalue les effets à long terme d'une action sur les récompenses cumulées, aidant à identifier la politique optimale π^* , qui maximise cette récompense attendue. La

fonction optimale Q^* , associée à π^* , s'exprime par (3.9).

$$Q^*(s_t, a_t) = \max_{\pi} \mathbb{E}[R_t \mid s_t = s, a_t = a] \quad (3.9)$$

Lorsque toutes les actions possibles a' pour l'état s' sont connues, l'action optimale maximise la récompense immédiate et future, comme exprimé par (3.10).

$$Q^*(s, a) = \mathbb{E} \left[r + \gamma \max_{a'} Q(s', a') \mid s_n = s, a_n = a \right] \quad (3.10)$$

Cette approche guide l'agent vers des actions maximisant les récompenses à long terme, en tenant compte des gains immédiats et actualisés.

DQN améliore le Q-learning en utilisant des DNN avec des paramètres θ pour approximer la fonction de valeur Q , $Q(s, a; \theta)$, dans des environnements à grande dimension. Le réseau de politiques guide l'agent en calculant les valeurs Q pour chaque paire état-action, tandis qu'un réseau cible stabilise les estimations de Q en évoluant plus lentement via des mises à jour différées.

Pour explorer la politique optimale, une stratégie ε -greedy est adoptée, équilibrant exploration et exploitation. Les actions sont sélectionnées aléatoirement avec une probabilité ε , tandis que les actions basées sur les valeurs Q sont choisies avec une probabilité $1 - \varepsilon$. Cela permet à l'agent d'améliorer progressivement ses décisions en naviguant dans l'espace des états.

Le DNN, noyau du Q-Network, est constitué de couches entièrement connectées avec des fonctions d'activation non linéaires comme ReLU (Rectified Linear Unit), exprimée

par (3.11).

$$f(z) = \max(0, z) \quad (3.11)$$

Ces couches capturent des relations complexes entre les paires état-action. La couche de sortie fournit les valeurs Q pour chaque action, guidant ainsi les décisions de l'agent.

Le DQN est formé à l'aide d'une fonction de perte qui minimise l'écart entre les valeurs Q prédites et les valeurs cibles, définies par (3.12).

$$y_t = r + \gamma \max_{a'} Q(s', a'; \theta') \quad (3.12)$$

où r est la récompense immédiate, γ le facteur d'actualisation, et $Q(s', a'; \theta')$ la valeur Q pour le prochain état. La fonction de perte s'exprime comme (3.13).

$$L(\theta) = \mathbb{E}_{s,a} \left[(y_t - Q(s, a; \theta))^2 \right] \quad (3.13)$$

Les paramètres θ sont mis à jour via la descente de gradient, comme indiqué dans (3.14).

$$\theta_{t+1} = \theta_t - \beta \nabla L(\theta_t) \quad (3.14)$$

Pour garantir la stabilité, un réseau cible est utilisé. Ses paramètres θ'_i évoluent progressivement selon (3.15).

$$\theta'_i \leftarrow \tau \theta_i + (1 - \tau) \theta'_i \quad (3.15)$$

où τ contrôle la lenteur de l'évolution.

Cette configuration, combinée à la relecture d'expérience où des transitions passées

(s_t, a_t, r_t, s_{t+1}) sont stockées, améliore la stabilité et réduit les corrélations dans les données. En conséquence, DQN peut approcher les valeurs Q optimales et accélérer l'apprentissage, tout en garantissant un équilibre efficace entre exploration et exploitation.

3.6.2.2 DQN pour l'optimisation RIS: une description algorithmique

Le cadre global de l'algorithme d'optimisation RIS est présenté dans la Fig.3.8. Cette approche basée sur DQN est bien adaptée pour optimiser les configurations RIS dans les systèmes de communication sans fil, où les espaces d'état et d'action sont vastes et complexes. Grâce à DQN, l'agent explore et exploite efficacement l'espace de configuration RIS, apprenant à maximiser la force du signal et à améliorer les performances globales de communication. Les éléments clés du cadre sont décrits ci-dessous:

- 1. Environnement:** l'environnement comprend l'émetteur, le RIS et le récepteur.
- 2. Agent:** Le RIS agit comme un agent.
- 3. État:** L'état s_t au moment t est défini dans (3.16) comme le RSSI actuel observé au Rx, noté RSSI_t .

$$s_t = \text{RSSI}_t \quad (3.16)$$

- 4. Action:** Une action a_t représente une configuration RIS spécifique au temps t . Le DQN sélectionne l'action optimale comme indiqué dans (3.17) en fonction de la valeur Q qui estime la récompense cumulée attendue.

$$a_t^* = \underset{a \in \mathcal{A}}{\operatorname{argmax}} Q(s_t, a) \quad (3.17)$$

Ici, \mathcal{A} est l'ensemble de toutes les actions possible. La valeur Q est calculée dans (3.18).

$$Q(s_t, a_t) = r_t + \gamma \cdot \max_{a' \in \mathcal{A}} Q(s_{t+1}, a') \quad (3.18)$$

où r_t est la récompense immédiate, γ est le facteur de remise et $Q(s_{t+1}, a')$ est la valeur Q estimée pour l'état suivant s_{t+1} .

Le DQN produit un vecteur de valeurs Q , noté $q = \pi(s_t)$, où chaque composante q_i correspond à la valeur estimée pour une action a_i . L'action optimale est alors sélectionnée comme le montre (3.19).

$$a_t^* = \underset{i}{\operatorname{argmax}} q_i, \quad \text{où } \mathbf{q} = \pi(s_t) \quad (3.19)$$

5. Récompense: La récompense r_t fournit un retour à l'agent, reflétant la qualité de son action a_t dans l'état donné. Il est calculé dans (3.20) comme un rapport normalisé, garantissant que les récompenses vont de 0 à 1, fournissant une mesure de performance claire pour évaluer les configurations du RIS.

$$r_t = \frac{\text{RSSI} - \text{minRSSI}}{\text{maxRSSI} - \text{minRSSI}} \quad (3.20)$$

L'algorithme 3 décrit les étapes impliquées dans l'optimisation des configurations RIS à l'aide du DQN, en tirant parti du DNN pour se rapprocher de la fonction de valeur Q $Q(s_t, a_t; \theta)$. Dans cette configuration, le RIS fonctionne aux côtés d'un émetteur et d'un récepteur pour améliorer les performances de communication sans fil. Le RSSI, appelé état s_t , mesuré au niveau du récepteur au pas de temps t , sert d'observation dans le modèle DQN. Le réseau DQN traite cet état et génère un ensemble de valeurs Q , $Q(s_t, a_t; \theta)$, qui

Algorithm 3 DQN pour l'Optimisation du RIS

- 1: **Entrées :** Taille du lot B , facteur d'actualisation γ , taux d'apprentissage β , taux d'exploration initial ϵ_{start} , taux d'exploration final ϵ_{end} , facteur de décroissance ϵ_{decay} , facteur de mise à jour douce τ , capacité de la mémoire de rappel C , nombre maximal d'étapes T , épisodes K .
- 2: **Sortie :** Configuration optimisée du RIS pour maximiser le RSSI.
- 3: Initialiser les paramètres du réseau de politique θ , les paramètres du réseau cible $\theta' \leftarrow \theta$, et la mémoire de rappel D .
- 4: **for** épisode $j = 1, \dots, K$ **do**
- 5: Réinitialiser l'environnement et obtenir l'état initial s_t .
- 6: **for** étape temporelle $t = 1, \dots, T$ **do**
- 7: Sélectionner une action a_t en utilisant une politique epsilon-greedy $\pi_\epsilon(s_t)$ avec un ϵ décroissant :

$$a_t = \begin{cases} \text{action aléatoire} & \text{avec une probabilité } \epsilon \\ \arg \max_a Q(s_t, a; \theta) & \text{avec une probabilité } 1 - \epsilon \end{cases} \text{ dans (3.17)}$$
- 8: Appliquer la configuration du RIS sélectionnée.
- 9: Mesurer le RSSI dans (3.16), observer l'état suivant s_{t+1} , et calculer la récompense r_t dans (3.20).
- 10: Stocker la transition (s_t, a_t, r_t, s_{t+1}) dans la mémoire de rappel D .
- 11: **if** taille de $D \geq B$ **then**
- 12: Échantillonner un mini-lot de transitions $\{s_j, a_j, r_j, s_{j+1}\}$ à partir de D .
- 13: Calculer la valeur cible y_j en utilisant (3.12)
- 14: Calculer la perte entre les valeurs Q prédites et les valeurs Q cibles en utilisant l'équation (3.7).
- 15: Effectuer une descente de gradient sur la perte $L(\theta)$ avec le taux d'apprentissage β en utilisant (3.13).
- 16: **end if**
- 17: Mettre à jour le réseau cible de manière douce en utilisant (3.15).
- 18: Mettre à jour l'état : $s_t \leftarrow s_{t+1}$.
- 19: **end for**
- 20: Réduire ϵ pour favoriser davantage l'exploitation au fil du temps :

$$\epsilon \leftarrow \epsilon_{\text{end}} + (\epsilon_{\text{start}} - \epsilon_{\text{end}}) \cdot \exp\left(-\frac{j}{\epsilon_{\text{decay}}}\right)$$

21: **end for**

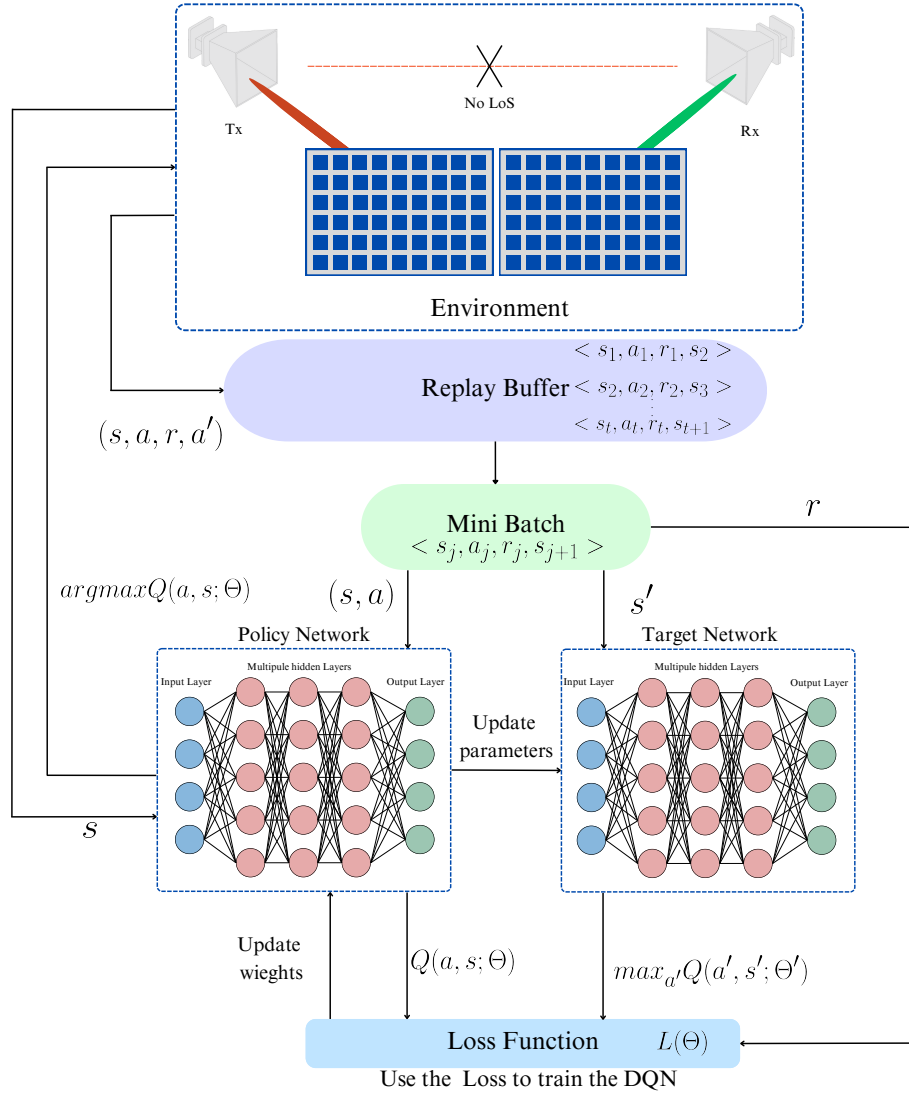


Figure 3.8 Architecture du RIS assistée par DQN

représentent les récompenses cumulées attendues pour différentes actions.

En se basant sur ces valeurs Q , le RIS sélectionne une action a_t selon une politique epsilon-greedy $\pi_\epsilon(s_t)$. L'action correspond à ajuster les déphasages des éléments RIS pour optimiser la réflexion du signal vers le récepteur. Après avoir pris l'action a_t , le système observe l'état suivant s_{t+1} , calcule la récompense r_t et passe à l'état suivant. L'interaction entre l'agent et l'environnement est capturée par le tuple de transition (s_t, a_t, r_t, s_{t+1}) , qui

est stocké dans un *replay buffer* D .

Une fois que le replay buffer contient suffisamment de transitions, le DQN échantillonne un mini-lot de transitions $\{(s_j, a_j, r_j, s_{j+1})\}$ de D pour mettre à jour le réseau de politique. Les valeurs Q cibles, $Q(s_{j+1}, a'; \theta')$, sont calculées à l'aide du réseau cible, où θ' représente les paramètres du réseau cible. La perte entre les valeurs Q prédites $Q(s_j, a_j; \theta)$ du réseau politique et les valeurs Q cibles est minimisée grâce à la descente de gradient.

Pour garantir la stabilité pendant la formation, les paramètres du réseau cible θ' sont mis à jour à l'aide d'un mécanisme de mise à jour logicielle, où les paramètres du réseau de politique θ sont fusionnés dans les paramètres du réseau cible θ' au fil du temps.

L'algorithme itère sur plusieurs épisodes, permettant à l'agent d'apprendre progressivement la configuration RIS optimale qui maximise les récompenses cumulées. La relecture d'expérience et les mises à jour logicielles assurent un apprentissage stable et une convergence vers une politique optimale.

3.6.2.3 DQN assisté par GA pour l'optimisation RIS: une description algorithmique

GA est reconnu pour sa capacité à relever des défis d'optimisation complexes, notamment dans les réseaux de communication sans fil. Cependant, DQN peut être limité lorsqu'il s'agit d'explorer de vastes espaces d'action ou de résoudre des problèmes hautement non linéaires, comme l'optimisation RIS. Une faiblesse notable du DQN réside dans son exploration aléatoire, qui peut entraîner une convergence lente et des performances sous-optimales. L'intégration de GA pallie ces limites en combinant l'exploration efficace de GA et le raffinement de DQN, offrant une solution prometteuse pour l'optimisation RIS.

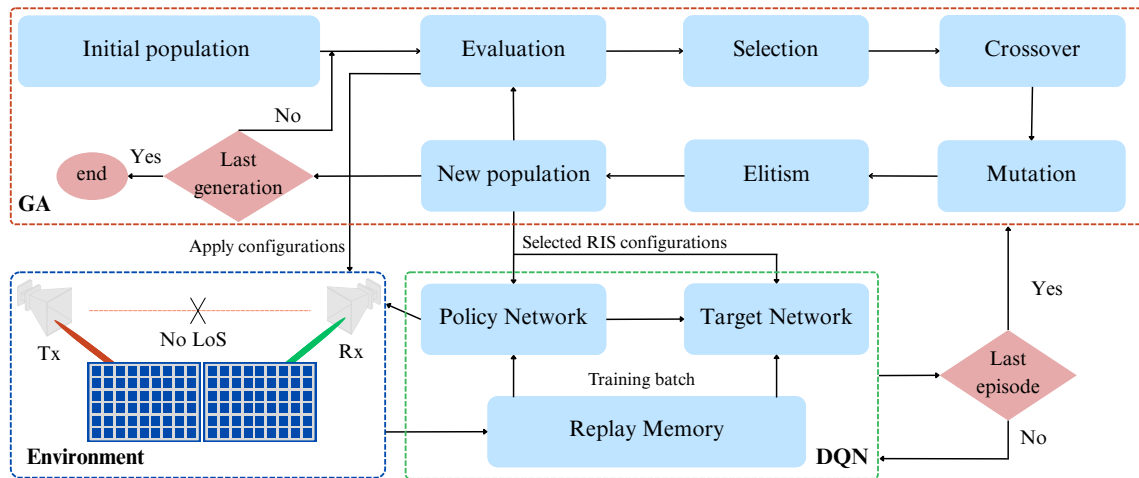


Figure 3.9 Architecture du DQN assistée par GA

La Fig.3.9 illustre l'architecture du DQN assisté par GA, où GA améliore les phases d'exploration et d'exploitation de DQN en fournissant des configurations de haute qualité. L'architecture montre un processus itératif où GA initialise l'espace d'action, ensuite affiné par DQN via son réseau politique.

Le processus débute par l'initialisation d'une population de configurations RIS. Chaque configuration est évaluée en mesurant le RSSI. Les meilleures configurations sont sélectionnées pour générer la prochaine génération via croisement et mutation, avec un mécanisme d'élitisme pour préserver les solutions les plus performantes.

Les configurations générées par GA sont ensuite affinées par DQN, qui établit une correspondance entre les configurations et leurs valeurs RSSI. Le réseau de politique DQN, soutenu par un réseau cible, apprend à prédire les meilleures configurations. Le réseau cible, mis à jour périodiquement, fournit des estimations stables, guidant les décisions en tenant compte des résultats à court et long termes.

Les configurations affinées sont testées dans l'environnement RIS, et les RSSI mesurés mettent à jour les réseaux de politique et cible du DQN. Un replay buffer stocke les

configurations et leurs RSSI, permettant d'apprendre des expériences passées tout en stabilisant le processus d'apprentissage.

Cette approche hybride, combinant l'exploration globale du GA et le raffinement du DQN, garantit une amélioration continue. Elle optimise efficacement les configurations RIS, démontrant un potentiel significatif pour améliorer l'optimisation du RSSI dans les systèmes de communication sans fil.

3.6.3 Bandit Adversaire

Dans le problème MAB (Multi-Armed Bandit), un décideur choisit parmi plusieurs options, chaque bras offrant une récompense inconnue suivant une distribution probabiliste. L'objectif est de maximiser la récompense cumulée tout en équilibrant exploration et exploitation. Cet équilibre est mesuré en termes de regret, défini comme la différence entre la récompense cumulée obtenue et celle qui aurait été atteinte en sélectionnant systématiquement le meilleur bras dès le début. [54,55].

Les formulations classiques de MAB supposent des distributions de récompenses stationnaires, adaptées à des environnements stables. Cependant, dans des contextes dynamiques comme l'optimisation du RIS pour les communications sans fil, où les conditions peuvent changer rapidement, ces hypothèses ne suffisent pas. Le cadre du AB devient pertinent : les récompenses varient arbitrairement dans le temps, sans hypothèses sur leur distribution. L'objectif est alors de minimiser le pire des cas, mesuré par le regret accumulé. Cette approche est particulièrement utile dans les systèmes RIS soumis à des conditions environnementales imprévisibles.

Pour de tels scénarios, l'algorithme EXP3 (Exponential-weight algorithm for Exploration and Exploitation) est souvent utilisé [56]. Il maintient des poids pour chaque

option, mis à jour en fonction des récompenses observées, et sélectionne les actions de manière probabiliste, équilibrant exploration et exploitation. Ce mécanisme rend EXP3 adapté aux environnements non stationnaires comme ceux des communications RIS.

L'algorithme EXP3 calcule la probabilité de choisir chaque action et met à jour les pondérations en fonction des récompenses. Les composants mathématiques incluent les formules de mise à jour des poids et de distribution des probabilités, démontrant l'efficacité d'EXP3 pour identifier les configurations optimales dans des environnements dynamiques. Ces étapes clés sont:

- **Distribution de probabilité des actions**

A chaque tour t , l'algorithme calcule la distribution de probabilité $P_i(t)$ sur les N actions possibles. Cette distribution de probabilité est exprimée dans (3.21).

$$P_i(t) = (1 - \gamma) \frac{w_i(t)}{\sum_{j=1}^N w_j(t)} + \frac{\gamma}{N}, \quad \text{for } i = 1, 2, \dots, N \quad (3.21)$$

Où:

- $w_i(t)$ est le poids de l'action $a_i(t)$ au tour t .
- γ est le paramètre d'exploration, contrôlant dans quelle mesure l'algorithme favorise l'exploration par rapport à l'exploitation, avec $\gamma \in [0, 1]$. Si γ est proche de 0, les poids domineront la sélection des actions favorisant ainsi l'exploitation, tandis qu'un γ proche de 1 augmentera l'exploration.
- N est le nombre total d'actions.

- **Sélection d'une action**

Une action $a_i(t)$ est sélectionnée selon la distribution de probabilité $P_i(t)$. L'action ayant la probabilité la plus élevée a plus de chance d'être choisie, mais l'exploration est assurée

par le terme γ , qui ajoute du hasard.

• Récompense et estimation

Après avoir sélectionné l'action $a_j(t)$, l'algorithme observe la récompense $r_j(t)$ pour cette action, où $r_j(t) \in [0, 1]$. Puisque les récompenses pour les autres actions ne sont pas observées, l'algorithme calcule une récompense estimée $\hat{r}_j(t)$ pour mettre à jour le poids de l'action. La récompense estimée est calculée en (3.22).

$$\hat{r}_j(t) = \begin{cases} \frac{r_j(t)}{P_j(t)} & \text{if } j = i_t \\ 0 & \text{otherwise} \end{cases} \quad \text{for } j = 1, \dots, N. \quad (3.22)$$

Cela garantit que la mise à jour est impartiale, même si seule la récompense de l'action sélectionnée est connue. Nous définissons le vecteur des récompenses estimées $\hat{R}(t)$ au tour t comme $\hat{R}(t) = (0, 0, \dots, 0, \hat{r}_j(t), 0, \dots, 0)$

• Mise à jour des poids

Une fois la récompense estimée $\hat{r}_j(t)$ calculée, le poids de chaque action est mis à jour de manière exponentielle. Le nouveau poids $w_j(t+1)$ pour l'action $a_j(t)$ est donné dans (3.23).

$$w_j(t+1) = w_j(t) \exp\left(\frac{\gamma \hat{r}_j(t)}{N}\right) \quad (3.23)$$

Cette mise à jour du poids garantit que les actions qui ont bien fonctionné (récompenses élevées) ont plus de chances d'être choisies lors des tours suivants.

• Regret cumulatif

Le regret dans le problème du bandit multi-armé mesure la différence entre la stratégie optimale et les performances de l'algorithme. Le regret $Reagret(t)$ au tour t est la différence entre la récompense cumulée de l'action optimale et la récompense cumulée

obtenue par l'algorithme selon (3.24).

$$\text{Regret}(t) = \sum_{s=1}^T r_{j^*}(s) - \sum_{s=1}^T r_j(s) \quad (3.24)$$

Où:

- $r_{j^*}(s)$ est la récompense de la meilleure action j^* au pas de temps s .
- et $r_j(s)$ est la récompense de l'action j choisie par l'algorithme au pas de temps s .
- T est le nombre de tours.

• Regret lié

Le regret lié au tour t représente la limite supérieure théorique du regret que l'algorithme peut accumuler. Cela donne l'assurance que, même dans le pire des cas (où les récompenses sont choisies de manière contradictoire), le regret ne dépassera pas cette limite. Le regret lié à l'algorithme EXP3 s'exprime dans (3.25).

$$\text{regretBound}(t) = (e - 1) \cdot \gamma \cdot \sum_{s=1}^T r_{j^*}(s) + \frac{N \cdot \log(N)}{\gamma} \quad (3.25)$$

où:

- e est le nombre d'Euler.

Dans le cas de l'optimisation du RIS, en appliquant l'algorithme EXP3, nous pouvons explorer de manière adaptative différentes configurations du RIS tout en exploitant celles qui maximisent systématiquement la qualité du signal. L'algorithme EXP3 pour l'optimisation RIS est représenté dans l'algorithme 4. À chaque itération, l'algorithme EXP3 calcule une distribution de probabilité pour les actions disponibles en fonction des poids actuels, qui sont initialisés de manière égale. Une action est sélectionnée en fonction de cette distribution de probabilité, et le RSSI pour la configuration choisie est

Algorithm 4 Algorithme EXP3 pour l'Optimisation du RIS

- 1: **Entrées** : Nombre d'actions N , paramètre d'exploration $\gamma \in [0, 1]$, nombre de tours T , configurations du RIS (actions).
 - 2: **Initialisation** : Initialiser les poids $w_i(1) = 1$ pour chaque action $a_i(1)$ avec $i \in \{1, \dots, N\}$
 - 3: **for** $t = 1, 2, \dots, T$ **do**
 - 4: Calculer la distribution de probabilité $P_i(t)$ pour les actions $a_i(t)$
 - 5: Sélectionner l'action $a_i(t)$ en fonction de la distribution de probabilité $P_i(t)$
 - 6: Recevoir la valeur du RSSI correspondant à l'action $a_i(t)$
 - 7: Calculer la récompense $r_i(t) \in [0, 1]$ pour l'action $a_i(t)$:

$$r_i(t) = \frac{\text{RSSI} - \text{minRSSI}}{\text{maxRSSI} - \text{minRSSI}}$$
 - 8: **for** $j = 1, 2, \dots, N$ **do**
 - 9: Estimer la récompense $\hat{r}_j(t)$
 - 10: Mettre à jour le poids pour l'action $a_j(t)$
 - 11: **end for**
 - 12: **end for**
 - 13: **Sortie** : La séquence d'actions choisies et leurs récompenses
-

mesuré. Le RSSI est ensuite normalisé pour calculer une récompense pour l'action. Cette récompense est utilisée pour estimer la véritable récompense, qui met à jour les poids de toutes les actions. L'algorithme vise à affiner continuellement les configurations RIS en privilégiant celles qui offrent de meilleures récompenses, tout en permettant l'exploration d'autres actions pour découvrir des configurations potentiellement meilleures. Cette approche permet une optimisation adaptative et dynamique des paramètres RIS dans un environnement non stationnaire.

3.6.4 AB assisté par GA

Nous introduisons l'intégration de GA avec le AB, en particulier l'algorithme EXP3, pour optimiser RIS. Bien que l'algorithme EXP3 fournisse une méthode robuste pour équilibrer l'exploration et l'exploitation dans des environnements non stationnaires, son efficacité peut être encore améliorée en tirant parti de la puissance évolutive des GA.

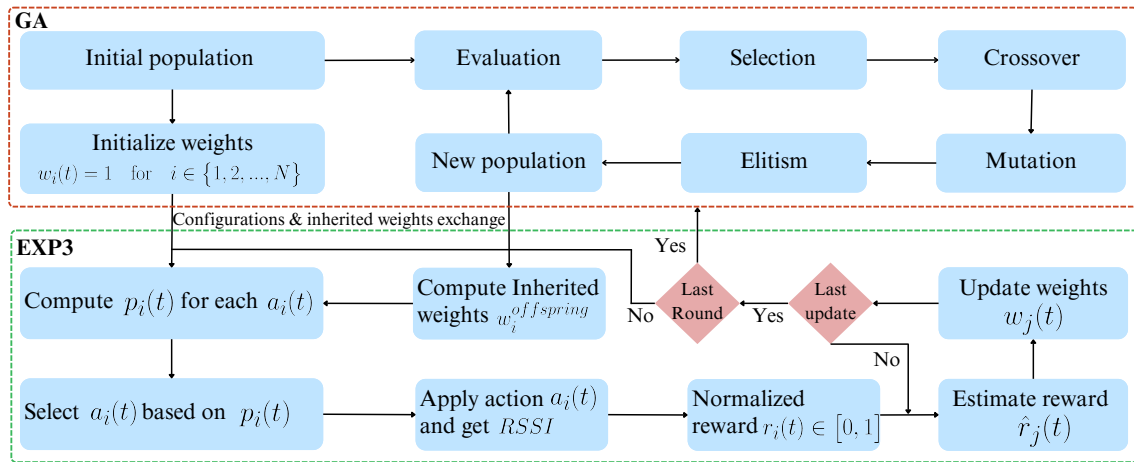


Figure 3.10 Architecture du AB assistée par GA

Les GA, connus pour leur capacité à explorer de grands espaces de configuration et à éviter les optima locaux, offrent un complément idéal à l'algorithme EXP3 en générant des configurations RIS optimisées basées sur des principes évolutifs. En combinant GA avec EXP3, l'approche AB assistée par GA qui en résulte permet au système de s'adapter plus efficacement aux environnements sans fil complexes et dynamiques. Cette méthode hybride permet une optimisation plus efficace, conduisant à une amélioration du RSSI et des performances globales du système.

La figure 3.10 illustre les étapes de mise en œuvre de l'algorithme AB assisté par GA. Le processus débute par l'initialisation d'une population aléatoire de configurations RIS de taille $N_{InitialPopulation}$, représentant différentes actions possibles. Chaque individu de la population est codé sous forme de vecteur binaire correspondant à une configuration RIS spécifique. Une fois la population initialisée, la phase d'évaluation commence. Après l'évaluation initiale de ces configurations en mesurant leur RSSI, l'algorithme EXP3 est appliqué à ces configurations RIS initiales sur un nombre prédéfini de tours. Pour la population initiale, le poids de chaque configuration, $w_i(1)$, est défini sur 1, ce qui signifie que toutes les configurations reçoivent une probabilité égale $p_i(1)$ au début. L'algorithme

EXP3 calcule ensuite la distribution de probabilité $p_i(t)$ pour chaque configuration RIS en fonction de leurs poids $w_i(t)$. Une action est sélectionnée en fonction de cette distribution de probabilité, la configuration est appliquée au RIS et le RSSI reçu est mesuré et une récompense $r_i(t)$ est calculée. Cette récompense est ensuite utilisée pour mettre à jour les poids des configurations. Celles qui fonctionnent bien reçoivent des pondérations plus élevées, tandis que les configurations les moins performantes reçoivent des pondérations plus faibles, ce qui permet à l'algorithme d'équilibrer entre l'exploration de nouvelles configurations et l'exploitation de celles qui produisent systématiquement une force de signal plus élevée. Une fois l'algorithme EXP3 appliqué et les premières itérations d'optimisation de la configuration du RIS terminées, les phases de sélection, croisement, mutation et évaluation du GA sont engagées pour créer la nouvelle génération. Ensuite, les poids de la nouvelle population sont mis à jour en héritant et en ajustant les poids calculés au cours des itérations EXP3. Cela permet aux configurations de conserver une mémoire de leurs performances tout en s'adaptant aux nouvelles configurations générées.

Dans le processus GA-assisté par AB, les poids de la génération précédente sont transmis aux descendants. Ce mécanisme d'héritage garantit que la performance de la génération précédente est conservée, mais ajustée pour refléter l'exploration introduite par le croisement. L'héritage des poids est déterminé en fonction de la probabilité de croisement, notée $p_{\text{crossover}}$. Pour chaque paire de parents, deux descendants sont générés, et les poids sont répartis entre eux comme suit :

1. Soit les poids des deux parents de la génération précédente w_{i-1}^{parent1} et w_{i-1}^{parent2} , respectivement.

2. Pour le premier descendant, le poids est calculé en (3.26).

$$w_i^{\text{offspring1}} = p_{\text{crossover}} \cdot w_{i-1}^{\text{parent1}} + (1 - p_{\text{crossover}}) \cdot w_{i-1}^{\text{parent2}} \quad (3.26)$$

Cela signifie que le poids du premier enfant est une moyenne pondérée du poids des parents, la proportion étant déterminée par la probabilité de croisement.

3. Pour le deuxième descendant, le poids est calculé de manière inverse en (3.27).

$$w_i^{\text{offspring2}} = p_{\text{crossover}} \cdot w_{i-1}^{\text{parent2}} + (1 - p_{\text{crossover}}) \cdot w_{i-1}^{\text{parent1}} \quad (3.27)$$

Cela garantit que les deux descendants reçoivent une combinaison du poids de leurs parents, mais dans des proportions différentes.

Une fois la nouvelle génération de descendants créée, l'algorithme EXP3 est appliqué à ces configurations. Contrairement à la génération initiale où les poids sont initialisés à 1, la nouvelle génération utilise les poids hérités de la génération précédente. Ces poids hérités représentent l'historique des performances des configurations parentes, garantissant que les connaissances issues des évaluations précédentes sont transmises aux descendants.

Les performances de chaque configuration sont utilisées pour mettre à jour les pondérations de manière dynamique. Comme dans la génération précédente, l'algorithme EXP3 calcule la distribution de probabilité de sélection de chaque configuration en fonction des poids hérités. Cependant, la principale différence est que les poids $w_i(1)$ ne sont plus initialisés à 1 mais sont directement hérités de la génération précédente basée sur le mécanisme de croisement. Ces poids guident le processus d'exploration et d'exploitation, permettant à l'algorithme de privilégier les configurations qui ont historiquement bien fonctionné, tout en continuant à explorer de nouvelles configurations

avec le paramètre d'exploration γ .

Après chaque tour, les poids sont mis à jour en fonction de la récompense observée et de la récompense estimée $\hat{r}_i(t)$. Ce processus itératif explore efficacement l'espace de configuration tout en exploitant les configurations réussies. La boucle de sélection, croisement, mutation, évaluation et application de l'algorithme EXP3 se répète sur un nombre défini de générations, affinant continuellement les configurations RIS pour atteindre des solutions optimales.

3.7 Conclusion

Dans ce chapitre, nous avons présenté les différents aspects méthodologiques qui sous-tendent cette étude. Nous avons tout d'abord détaillé les composants matériels utilisés, suivi par une explication du principe de fonctionnement du RIS et de l'intégration des divers éléments dans le système global. Enfin, nous avons décrit les approches adoptées pour optimiser les configurations du RIS, mettant en lumière les algorithmes et procédures déployés.

Ces éléments méthodologiques constituent la base nécessaire pour comprendre les résultats et analyses présentés dans le chapitre suivant. Nous examinerons comment les approches décrites ici ont permis d'atteindre les objectifs de l'étude et d'optimiser les performances du système dans des conditions expérimentales.

Chapitre 4 - Expérimentation et résultats

Dans ce chapitre, une installation expérimentale est mise en place pour évaluer les performances des algorithmes d'optimisation implémentés. À partir des résultats obtenus dans ce cadre, nous analysons l'efficacité de chaque méthode dans des conditions réelles, définies précisément pour ce projet. Cette analyse vise à mesurer la pertinence des approches proposées pour l'optimisation du RIS, tout en mettant en lumière leur comportement en environnement pratique.

4.1 Originalité de l'approche expérimentale

Il est important de préciser que certaines composantes de l'installation expérimentale s'inspirent de dispositifs déjà décrits dans la littérature, notamment en ce qui concerne le matériel utilisé. Toutefois, la majorité des travaux existants sur l'optimisation des RIS sont réalisés par simulation, sans validation expérimentale en conditions réelles. Ce mémoire se distingue par une approche pratique et expérimentale, reposant sur des tests en temps réel avec une RIS conçue sur mesure. L'ensemble des algorithmes d'optimisation a été implémenté, testé et comparé à travers une plateforme réelle. Cette démarche expérimentale constitue l'une des principales originalités du projet, offrant une contribution concrète à l'évaluation des performances des RIS dans des scénarios applicatifs proches de la réalité.

4.2 Description de l'installation expérimentale

La configuration utilisée est représentée dans la Fig.4.3. Deux radios définies par logiciel (SDR) ADALM-PLUTO ont été employées. Ces dispositifs ont été configurés lors d'un projet antérieur mené au sein du LSSI (*voir section 2.5*) portant sur le déploiement de plateformes SDR, l'une pour l'émission et l'autre pour la réception. Les antennes cornet,

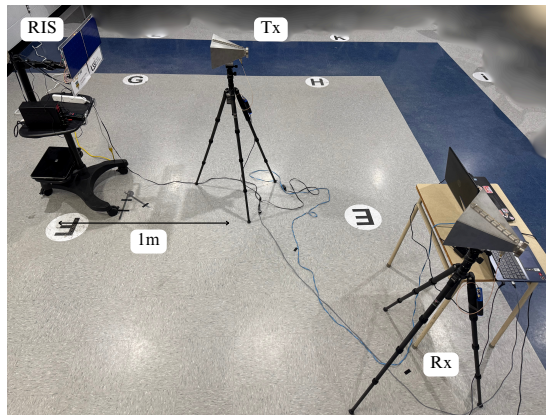


Figure 4.1 Vue de dessus de l'installation

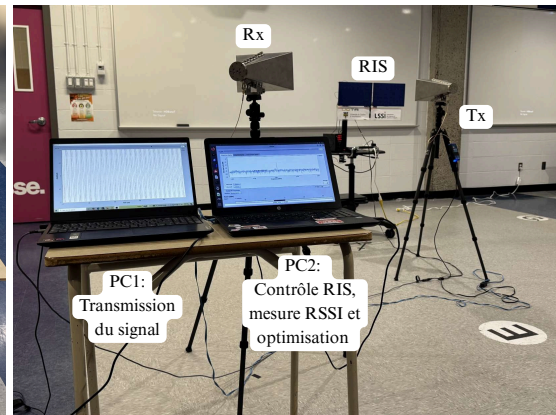


Figure 4.2 Vue latérale de l'installation

Figure 4.3 Configuration avec différents points de vue

utilisées pour transmettre et recevoir, assurent une focalisation directionnelle du signal, évitant ainsi toute LoS.

Deux RIS, fonctionnant à 3,75 GHz, ont été intégrées, chacune comportant 48 éléments. Leur conception matérielle ainsi que leur configuration logicielle constituent une composante centrale de ce projet et sont détaillées dans la *section 3.4*. Contrôlées par un même microcontrôleur, ces cartes fonctionnent comme une seule surface réfléchissante composée de 96 éléments, élargissant ainsi la zone de réflexion.

4.2.1 Positionnement des équipements

Les tests ont été effectués dans une salle, où le RIS est placé au centre. Le transmetteur est positionné directement devant le centre du RIS à une distance de 1 m et à un angle de 90° . Le récepteur est placé à différents angles et distances par rapport au centre du RIS pour mesurer la performance dans plusieurs configurations. Les positions spécifiques du récepteur sont décrites dans le tableau 4-1 et illustrées dans la Fig.4.4.

Tableau 4-1 Positions du récepteur par rapport au RIS

Position	Angle (°)	Distance (m)
Pos1	45	3
Pos2	60	1.5
Pos3	75	2.5
Pos4	105	1.5
Pos5	120	2.5
Pos6	135	2

4.2.2 Génération du signal

Le transmetteur produit un signal modulé en onde continue, une onde cosinusoidale. Le script de transmission génère une fréquence de tonalité de 50 kHz avec un taux d'échantillonnage de 2,084 MHz. Ce signal est transmis à une fréquence centrale de 3,75 GHz avec une puissance de sortie de -30 dBm. L'antenne cornet utilisée pour le transmetteur possède un gain isotrope de 11 dBm à 3,75 GHz, focalisant et amplifiant le signal dans une direction spécifique. Ce gain permet une augmentation de puissance de 11 dBm, améliorant ainsi la réception du signal réfléchi par le RIS. Cette chaîne d'émission est implémentée dans GNU Radio sur un PC dédié, qui pilote un module PlutoSDR configuré pour émettre en continu. Le signal sinusoïdal est généré numériquement, converti en signal RF par le PlutoSDR, puis transmis via l'antenne. Une interface graphique permet de surveiller en temps réel la forme d'onde transmise, assurant ainsi la stabilité et la cohérence du signal durant toute l'expérimentation.

4.2.3 Réception et optimisation

Du côté de la réception, une antenne directionnelle capte le signal réfléchi par le RIS à différents emplacements. Le module PlutoSDR utilisé en réception est configuré à une fréquence centrale de 3,75 GHz avec un gain réglé manuellement à 64 dB et un taux d'échantillonnage de 2,084 MHz. Le signal reçu, sous forme de flux I/Q, est traité pour en extraire la puissance moyenne en temps réel.

Le Rx mesure le RSSI pour fournir un retour d'information au système de contrôle RIS. Ce processus de rétroaction en temps réel consiste à :

1. Capturer le signal reçu à partir du flux RF.
2. Estimer sa puissance moyenne (RMS), convertir cette valeur en dBm, puis en calculer la moyenne glissante.
3. Transmettre la valeur mesurée au système de contrôle RIS, qui l'utilise pour évaluer l'efficacité de la configuration courante.

La valeur du RSSI est actualisée toutes les 100 millisecondes et affichée en temps réel, garantissant un retour rapide et continu vers le système d'optimisation. Cette boucle de rétroaction dynamique permet d'ajuster les configurations du RIS à chaque itération, assurant ainsi une amélioration optimale du signal et une couverture ciblée dans la direction souhaitée.

4.3 Comparaison des performances des méthodes d'optimisation RIS dans les configurations spatiales

Dans cette section, nous analysons les performances des différentes méthodes d'optimisation RIS dans les configurations spatiales spécifiques, en tenant compte de leurs paramètres d'implémentation. Le tableau 4-2 résume les principaux paramètres utilisés

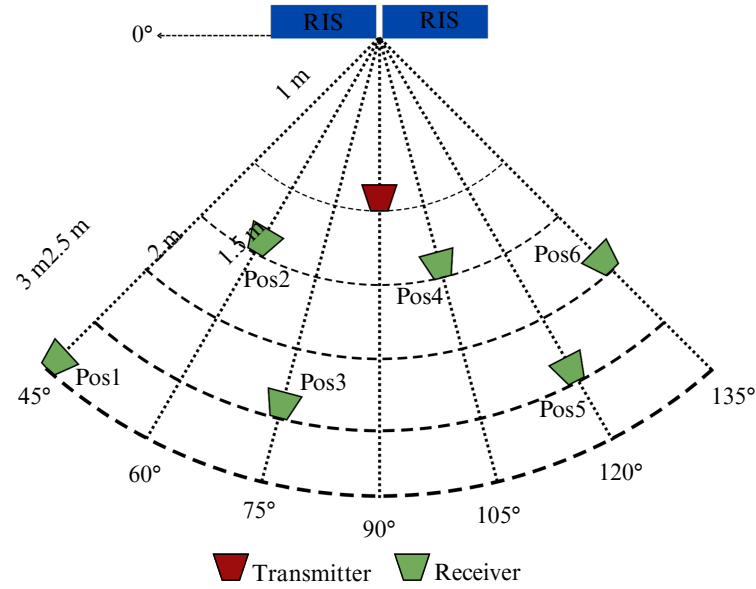


Figure 4.4 Positions de test de Pos1 à Pos6

pour chaque méthode d'optimisation étudiée. Les paramètres indiqués dans le tableau ont été définis une seule fois et appliqués de manière identique à l'ensemble des positions de test (Pos1 à Pos6) présentées dans le tableau 4-1. Cela permet une comparaison équitable des performances des différentes méthodes, indépendamment des variations spatiales du récepteur. Les paramètres de chaque méthode ont été sélectionnés pour optimiser leurs performances dans des environnements sans fil réels, en équilibrant l'exploration des configurations RIS et leur exploitation. Par exemple, le taux de mutation (p_{mutation}) et le taux de croisement ($p_{\text{crossover}}$) du GA favorisent une exploration approfondie tout en maintenant des solutions prometteuses, tandis que le facteur γ dans EXP3 contrôle le compromis entre exploration et exploitation. De même, le taux d'apprentissage (LR , Learning Rate) et les paramètres d'épsilon dans le DQN influencent la vitesse et la stabilité de la convergence.

Tableau 4-2 Tableau des paramètres pour les algorithmes GA, DQN et EXP3

GA		DQN		EXP3	
Name	Value	Name	Value	Name	Value
$N_{InitialPopulation}$	256	BATCH_SIZE	128	numActions	$len(population)$
$N_{Generations}$	200	GAMMA	0.99	γ	0.07
$N_{Parents}$	130	EPS_START	0.9	Weights (initial)	1.0
$N_{Solutions}$	256	EPS_END	0.05	Rounds	50
N_{Elites}	20	EPS_DECAY	1000	-	-
$p_{crossover}$	0.9	TAU	0.005	-	-
$p_{mutation}$	0.3	LR	1×10^{-4}	-	-

Les performances des algorithmes ont ensuite été comparées en termes de gains de RSSI, comme le montre la Fig. 4.5, et leur capacité à s'adapter aux différentes positions du récepteur, illustrée par la Fig. 4.6. Ces résultats démontrent l'impact direct de ces paramètres sur l'efficacité des méthodes étudiées.

La Fig.4.5 et la Fig.4.6 illustrent les performances des méthodes d'optimisation RIS sur six positions de récepteur, définies par des angles et des distances spécifiques dans le tableau4-1. La Fig.4.5, représentée sous forme de graphique à barres, met en évidence la force du signal soutenue obtenue par chaque méthode, basée sur le CMA RSSI (Cumulative Moving Average du Received Signal Strength Indicator). Il est important de noter que les valeurs minimales observées sont identiques pour toutes les méthodes ; elles correspondent à la valeur initiale du RSSI mesurée avant l'application de toute configuration optimisée. Ce point de départ commun permet une évaluation comparative équitable des performances des différentes approches. En revanche, la Fig.4.6, sous forme

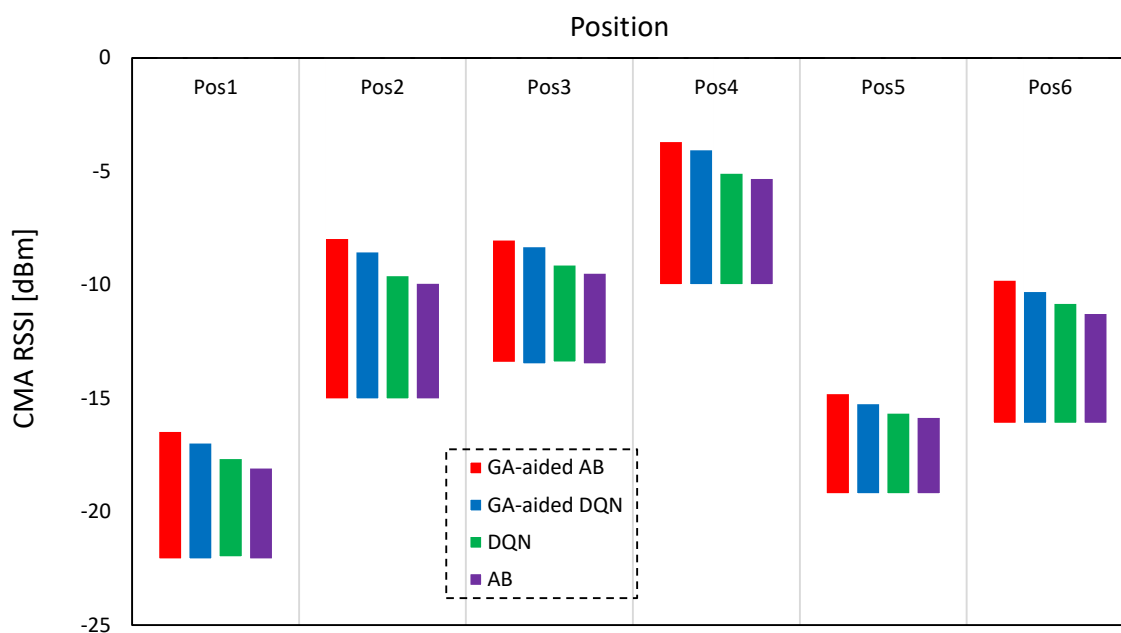


Figure 4.5 Diagramme en barres comparant les valeurs CMA RSSI pour diverses méthodes d'optimisation à plusieurs positions de récepteur

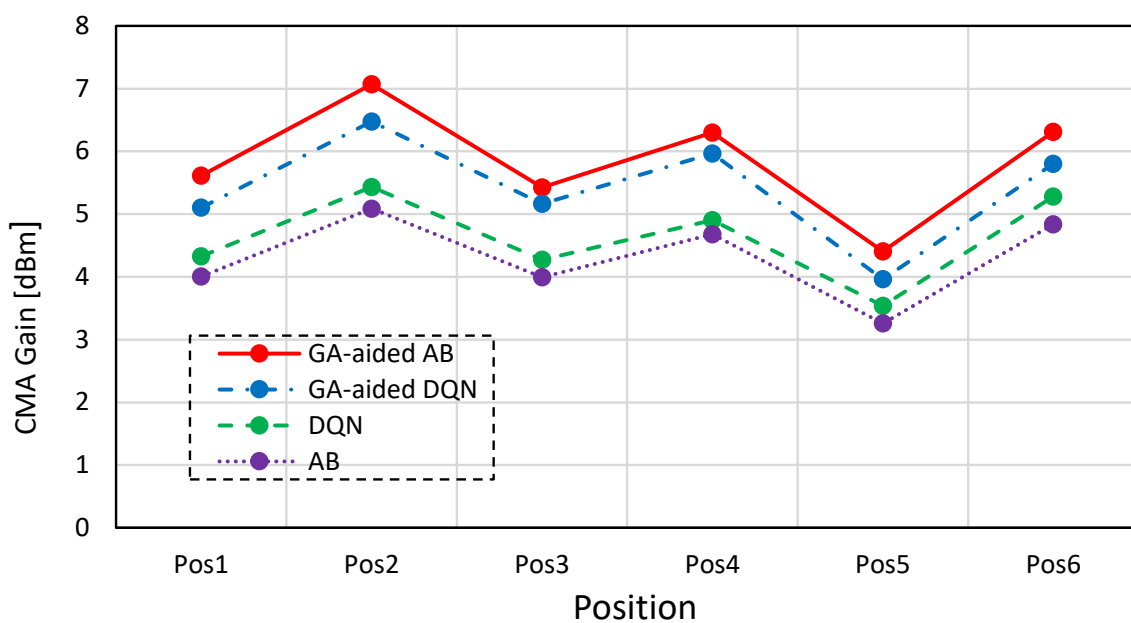


Figure 4.6 Performance du gain CMA des méthodes d'optimisation RIS à différentes positions

Tableau 4-3 Comparaison des performances des algorithmes d'optimisation RIS avec valeurs minimales et maximales des gains

Critères	AB assisté par GA	DQN assisté par GA	DQN (autonome)	AB (autonome)
Gain minimum (dBm)	4,41	3,96	3,54	3,26
Gain maximum (dBm)	7,07	6,47	5,43	5,09
Convergence	Exploration approfondie	Convergence rapide	Convergence lente	Convergence lente
	Stabilisation lente	Exploitation efficace	Exploration limitée	Faible adaptation
Adaptabilité	Robuste toutes à positions	Bonne dans Pos2 et Pos4	Limité dans Pos5	Faible en conditions complexes
	Excellente en Pos5	Performant en Pos2	Faible exploration	Moins performant

de graphique linéaire, présente l'évolution du CMA gain, défini comme la différence entre le RSSI maximal et le RSSI minimal au fil des itérations. Ces deux représentations offrent une vision complémentaire des performances des méthodes, mettant en lumière la tendance, la convergence et les variations des résultats au cours de l'optimisation

4.3.1 Performances générales des algorithmes d'optimisation RIS

Les résultats démontrent que l'AB assisté par GA (GA-aided AB) se distingue systématiquement comme la méthode d'optimisation la plus performante dans toutes les positions du récepteur, atteignant les valeurs CMA RSSI les plus élevées. Comme le résume le tableau 4-3, l'AB assisté par GA réalise des gains impressionnants de 7,1 dBm à Pos2 (60°, 1,5 m) et de 6,3 dBm à Pos6 (135°, 2 m), illustrant sa capacité à explorer et affiner efficacement les configurations RIS. Cette performance est attribuable à la complémentarité entre les larges capacités d'exploration du GA et le mécanisme de

renforcement d'AB, qui permet à cette méthode de naviguer dans un vaste espace de solutions tout en optimisant les configurations pour maximiser la force du signal.

Le DQN assisté par GA affiche également d'excellentes performances, avec des gains notables de 6,5 dBm à Pos2 et de 5,8 dBm à Pos6. Bien qu'il reste légèrement en deçà des performances maximales de l'AB assisté par GA, cette méthode bénéficie de l'intégration de GA, qui améliore significativement ses capacités d'exploration et d'optimisation. La stabilisation rapide du DQN assisté par GA le rend particulièrement efficace dans des configurations où l'équilibre entre exploration et exploitation est essentiel.

En revanche, les méthodes autonomes comme DQN et AB montrent des limites importantes pour maximiser le RSSI. Le DQN atteint un gain maximum de 5,4 dBm, surpassant légèrement l'AB autonome, qui obtient un gain maximum de 5,1 dBm. Cependant, ces performances restent nettement inférieures à celles des méthodes assistées par GA, mettant en évidence le rôle crucial de GA dans l'amélioration des performances d'optimisation.

Ces résultats soulignent l'importance des approches hybrides qui combinent exploration approfondie et mécanismes de renforcement, telles que l'AB assisté par GA, pour relever les défis de l'optimisation RIS dans des environnements complexes.

4.3.2 Impact des positions spatiales sur les performances des méthodes d'optimisation RIS

Les résultats soulignent l'influence cruciale de la position du récepteur sur les performances des différentes méthodes d'optimisation RIS. Les positions présentant des angles modérés et des distances plus courtes, telles que Pos2 (60°, 1,5 m) et Pos4 (105°, 1,5 m), favorisent les performances optimales des méthodes assistées par GA. Par exemple,

l'AB assisté par GA atteint des gains maximaux de 7,1 dBm à Pos2 et de 6,3 dBm à Pos4, démontrant sa capacité à exploiter efficacement les configurations RIS pour maximiser le RSSI dans ces scénarios. De manière similaire, le DQN assisté par GA affiche d'excellents résultats à ces positions, atteignant 6,5 dBm à Pos2 et 6,0 dBm à Pos4, bien que sa convergence plus rapide limite parfois son exploration comparée à l'AB assisté par GA.

Les positions modérées, comme Pos1 (45° , 3 m) et Pos6 (135° , 2 m), offrent des performances équilibrées. À Pos1, l'AB assisté par GA atteint un gain significatif de 5,6 dBm, tandis que le DQN assisté par GA réalise un gain de 5,1 dBm, illustrant leur capacité à maintenir une optimisation robuste même à des distances légèrement plus longues. À Pos6, l'AB assisté par GA conserve un net avantage avec un gain de 6,3 dBm, suivi de près par le DQN assisté par GA, qui atteint 5,8 dBm.

En revanche, les positions défavorables telles que Pos5 (120° , 2,5 m), caractérisées par des angles plus larges et des distances accrues, mettent en évidence les limites des méthodes d'optimisation autonomes. Malgré ces défis, l'AB assisté par GA affiche une performance supérieure, atteignant un gain de 4,4 dBm, surpassant le DQN (3,5 dBm) et l'AB autonome (3,3 dBm), démontrant son adaptabilité à des conditions moins favorables.

Des positions comme Pos3 (75° , 2,5 m) représentent des scénarios intermédiaires. Dans cette configuration, l'AB assisté par GA atteint un gain de 5,4 dBm, tandis que le DQN assisté par GA enregistre 5,2 dBm, montrant une légère baisse par rapport aux positions optimales. Les méthodes autonomes, quant à elles, peinent à dépasser les 4,0 dBm dans ces conditions.

Ces résultats mettent en évidence que l'efficacité des méthodes d'optimisation est fortement influencée par les caractéristiques spatiales, notamment les angles et les

distances. Les approches assistées par GA, en particulier l'AB assisté par GA, se distinguent par leur robustesse et leur capacité à s'adapter à une large gamme de configurations, maximisant systématiquement les gains de RSSI même dans des scénarios complexes.

4.4 Comportement de convergence des méthodes d'optimisation assistées par GA et autonomes

Dans la Fig.4.7, nous affichons les courbes de convergence des valeurs CMA RSSI pour quatre méthodes d'optimisation RIS différentes : AB assisté par GA, DQN assisté par GA, DQN et AB, sur six positions de récepteur. Chaque courbe illustre le CMA RSSI à mesure qu'il se stabilise au fil du temps, fournissant une représentation visuelle des performances de chaque méthode à mesure qu'elle s'approche d'une configuration optimale. Toutes les méthodes ont été testées sur un nombre identique d'itérations, avec un volume de données constant à chaque position, ce qui garantit une base de comparaison équitable entre les courbes et permet d'attribuer les différences observées aux performances réelles des algorithmes.

Dans toutes les positions, l'AB assisté par GA atteint systématiquement les valeurs RSSI CMA les plus élevées. Sa courbe démontre une stabilisation plus lente par rapport aux autres méthodes, indiquant une exploration soutenue des configurations RIS sur plus d'itérations. Cette exploration étendue permet à AB assisté par GA d'améliorer progressivement le RSSI et d'atteindre des valeurs plus élevées que les méthodes autonomes et même le DQN assisté par GA ne peuvent pas atteindre. Ce comportement reflète l'effet combiné de l'exploration approfondie de GA, qui élargit l'espace de recherche, et du mécanisme de renforcement d'AB, qui affine les configurations pour maximiser la force du signal. La capacité de l'AB assisté par GA à s'adapter à

différentes configurations spatiales met en évidence sa robustesse, en particulier dans des positions difficiles telles que Pos5, où des angles et des distances sous-optimaux entravent considérablement les autres méthodes.

Le DQN assisté par GA fonctionne également bien, atteignant les deuxièmes valeurs RSSI les plus élevées parmi les postes. Cependant, le DQN assisté par GA se stabilise plus rapidement que l'AB assisté par GA, ce qui suggère une exploration moins intensive. Cette stabilisation plus rapide indique que DQN assisté par GA exploite efficacement les configurations prometteuses, mais pourrait manquer les configurations les plus performantes découvertes par l'exploration prolongée de AB assisté par GA. Par conséquent, bien que DQN assisté par GA bénéficie des améliorations apportées par GA, il n'explore pas l'espace de configuration de manière approfondie que AB assisté par GA. Ses performances sont particulièrement fortes dans les positions favorables, telles que Pos2 et Pos4, où les configurations spatiales s'alignent bien avec sa stratégie d'optimisation.

Les DQN et AB autonomes affichent une convergence plus lente et des valeurs RSSI finales inférieures. Parmi celles-ci, DQN surpasse généralement AB, mais les deux méthodes plafonnent à des niveaux RSSI nettement inférieurs à ceux de leurs homologues assistés par GA. Cela suggère que sans les capacités d'exploration de GA, les méthodes autonomes sont moins efficaces pour naviguer dans l'espace de configuration RIS pour découvrir des solutions optimales. De plus, leurs performances limitées sont plus prononcées dans les positions difficiles, telles que Pos5, où des chemins de réflexion sous-optimaux limitent davantage l'amélioration du signal. Dans l'ensemble, les courbes de convergence mettent en évidence la supériorité des méthodes assistées par GA, en particulier l'AB assistée par GA, qui équilibre efficacement l'exploration et l'exploitation

pour atteindre les valeurs RSSI les plus élevées dans diverses configurations spatiales. Le DQN assisté par GA offre une alternative compétitive avec une stabilisation plus rapide. Les méthodes autonomes sont considérablement à la traîne, soulignant le rôle essentiel de GA dans l'amélioration de l'optimisation des RIS.

4.5 Synthèse comparative des performances des algorithmes

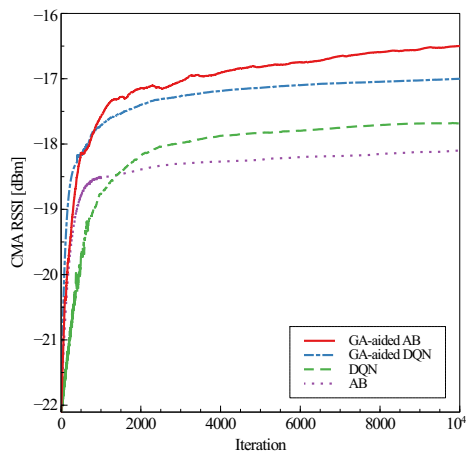
Les performances des méthodes d'optimisation RIS ont été évaluées selon plusieurs critères qualitatifs : robustesse, convergence, adaptabilité, complexité et utilisation des ressources. Le tableau 4-4 propose une synthèse comparative de ces algorithmes, mettant en évidence leurs points forts et leurs limites dans différents contextes expérimentaux.

Ces critères sont définis en fonction de leur rôle dans l'évaluation des performances. La convergence correspond à la rapidité et à la stabilité avec lesquelles chaque méthode atteint une performance optimale. L'adaptabilité désigne la capacité d'un algorithme à maintenir des performances élevées malgré les variations des positions du récepteur. La complexité englobe à la fois la charge algorithmique, incluant le nombre d'opérations et la structure de traitement, et les efforts liés à la mise en œuvre, tels que la programmation, les réglages et la stabilité du système. Les ressources concernent l'utilisation du processeur et de la mémoire vive, observées de manière qualitative à travers le comportement du système. Les évaluations reposent sur des observations empiriques réalisées lors des expérimentations en temps réel, notamment à travers l'analyse des courbes de convergence et des performances mesurées sur la plateforme de test.

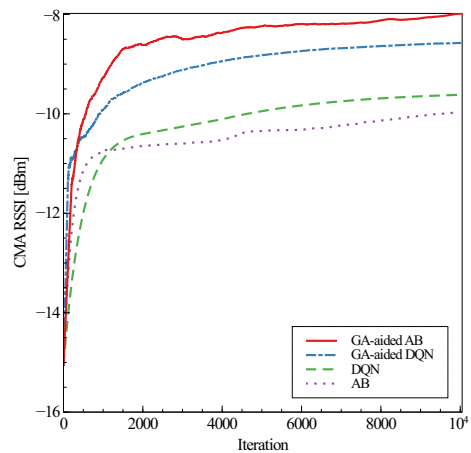
4.6 Conclusion

Les résultats obtenus mettent en évidence les performances supérieures des méthodes d'optimisation assistées par le GA, en particulier l'AB assisté par GA. Ces approches

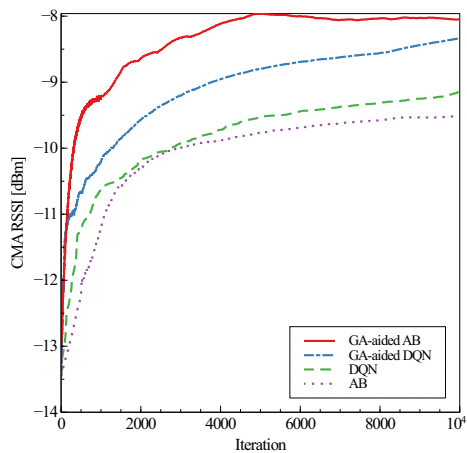
montrent une capacité robuste à explorer efficacement l'espace de configuration du RIS, même dans des conditions difficiles. Ces résultats soulignent également le rôle crucial de l'équilibre entre exploration et exploitation dans l'optimisation des performances du système. Les conclusions obtenues fournissent une base solide pour de futures investigations sur des scénarios encore plus complexes, tels que les environnements multi-utilisateurs ou les configurations multi-RIS.



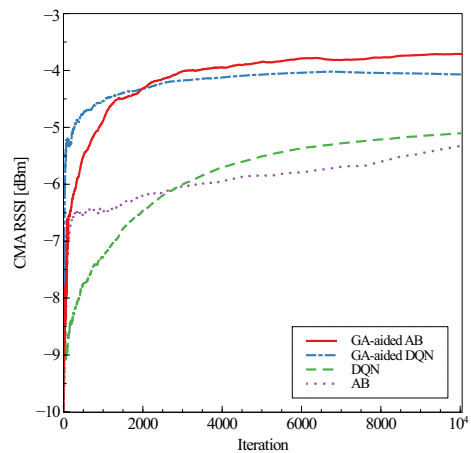
(a) Pos1



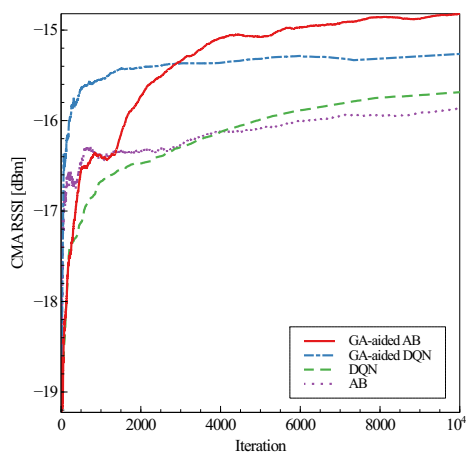
(b) Pos2



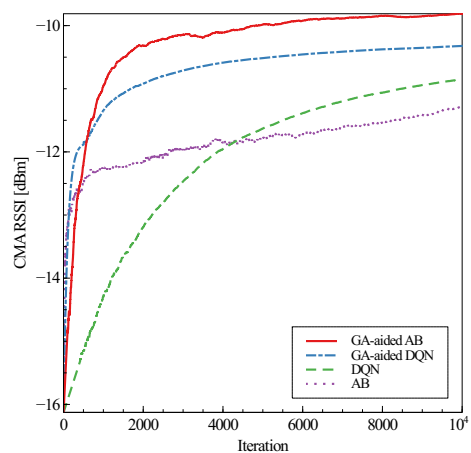
(c) Pos3



(d) Pos4



(e) Pos5



(f) Pos6

Figure 4.7 Courbes de convergence des algorithmes d'optimisation pour six positions

Tableau 4-4 Résumé des Performances des Algorithmes d'Optimisation RIS

Critères	AB assisté par GA	DQN assisté par GA	DQN (autonome)	AB (autonome)
Exploration	Étendue, grâce aux capacités du GA, couvrant un large espace	Assez large mais sujette à une stabilisation précoce	Limitée à des configurations locales	Réduite, principalement orientée vers l'exploitation
Convergence	Lente mais aboutissant à des améliorations continues	Rapide, avec un risque de convergence prématurée	Modérée, aboutissant à une convergence partielle	Plus lente et souvent sous-optimale
Adaptabilité	Excellente, y compris dans des scénarios complexes	Élevée dans des conditions favorables, mais sensible aux changements	Moyenne, limitée en présence d'obstacles	Restreinte, se révèle peu performante dans des environnements variés
Performances globales	Excellentes, conciliant exploration approfondie et exploitation efficace	Solides, bien que légèrement en deçà d'AB assisté par GA	Acceptables, notamment dans des configurations simples	Modestes, fréquemment entravées par les limites du modèle
Complexité	Modérée, avec la nécessité de paramétrer le GA	Intermédiaire, requiert une intégration fluide du GA	Faible, mais accompagnée de performances restreintes	Très faible, mise en œuvre aisée mais efficacité limitée
Utilisation des ressources	Optimisée, proposant un compromis satisfaisant entre temps et calcul	Correcte, avec un rendement appréciable malgré une stabilité parfois délicate	Moyenne, requiert des ajustements fréquents	Faible, peu exigeante mais peu efficace
Robustesse	Très robuste face aux variations spatiales et conditions défavorables	Robuste dans des scénarios ciblés	Moyennement robuste, vulnérable aux positions extrêmes	Fragile dans les environnements dynamiques, limitant ses performances

Chapitre 5 - Conclusion

Le projet présenté s'inscrit dans le cadre des avancées technologiques visant à transformer les communications sans fil par l'intégration de RIS. L'objectif principal de cette étude était de contrôler et optimiser un RIS personnalisé, en explorant ses configurations à l'aide d'approches algorithmiques avancées telles que les GA, le DQN et AB. En particulier, l'algorithme AB assisté par GA s'est révélé être une solution performante, démontrant des capacités exceptionnelles dans des environnements contrôlés et des scénarios complexes.

L'un des aspects centraux de ce projet a été le développement d'un contrôle précis et une optimisation flexible du RIS personnalisé, composé d'éléments configurables. Ce contrôle a permis une navigation efficace dans l'espace des solutions, facilitant l'identification des configurations optimales pour maximiser le RSSI. Grâce à l'intégration d'approches hybrides combinant les GA et les AB, le système a démontré une capacité sans précédent à adapter dynamiquement les configurations du RIS en réponse à des variations environnementales. Les résultats expérimentaux, validés sur six positions spécifiques du récepteur dans un environnement intérieur, confirment que l'approche AB assisté par GA excelle par rapport aux méthodes traditionnelles. En maximisant systématiquement le RSSI, elle a démontré une capacité exceptionnelle à naviguer efficacement dans l'espace des solutions pour identifier les configurations RIS optimales.

L'interaction entre exploration et exploitation, permise par l'intégration des GA avec les mécanismes des AB, constitue une avancée significative. Cette combinaison a permis d'élargir l'espace de recherche tout en affinant les configurations pour maximiser la force du signal. Toutefois, les méthodes autonomes, bien qu'elles montrent une certaine efficacité, restent limitées en raison de leur incapacité à équilibrer efficacement ces deux

aspects fondamentaux de l'optimisation. Par ailleurs, l'étude a mis en lumière des défis liés à l'extension des approches actuelles à des environnements plus complexes, notamment ceux impliquant des scénarios multi-utilisateurs ou multi-RIS.

En définitive, cette étude a permis de valider les hypothèses formulées au début du projet, mais elle a également ouvert de nouvelles pistes de réflexion sur les défis dans l'intégration des RIS dans les réseaux de communication de nouvelle génération.

Références

- [1] K. David and H. Berndt, “6g vision and requirements: Is there any need for beyond 5g?” *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 72–80, 2018.
- [2] M. Sauter, *Evolution from 2G over 3G to 4G*, 2013, pp. 1–7.
- [3] H. Lee, B. Lee, H. Yang, J. Kim, S. Kim, W. Shin, B. Shim, and H. V. Poor, “Towards 6g hyper-connectivity: Vision, challenges, and key enabling technologies,” *Journal of Communications and Networks*, vol. 25, no. 3, pp. 344–354, 2023.
- [4] A. I. Salameh and M. El Tarhuni, “From 5g to 6g—challenges, technologies, and applications,” *Future Internet*, vol. 14, no. 4, p. 117, 2022.
- [5] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, “Small-scale, local area, and transitional millimeter wave propagation for 5g communications,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6474–6490, 2017.
- [6] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, “5g: A tutorial overview of standards, trials, challenges, deployment, and practice,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.
- [7] M. Z. Siddiqi and T. Mir, “Reconfigurable intelligent surface-aided wireless communications: An overview,” *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 33–63, 2022.
- [8] J. An, M. Debbah, T. J. Cui, Z. N. Chen, and C. Yuen, “Emerging technologies in intelligent metasurfaces: Shaping the future of wireless communications,” *arXiv preprint arXiv:2411.19754*, 2024.
- [9] M. D. Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, “Smart radio environments empowered by reconfigurable ai meta-surfaces: An idea whose time has come,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, 2019.
- [10] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, “Reconfigurable intelligent surfaces: Principles and opportunities,” *IEEE Communications Surveys Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.
- [11] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elrashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, “Reconfigurable intelligent surfaces for 6g systems: Principles, applications, and research directions,” *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, 2021.
- [12] 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, no. 5, pp. 3313–3351, 2021.
- [13] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.

- [14] A. Welkie, L. Shangguan, J. Gummesson, W. Hu, and K. Jamieson, "Programmable radio environments for smart spaces," in *Proceedings of the 16th ACM Workshop on Hot Topics in Networks*, 2017, pp. 36–42.
- [15] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "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, no. 11, pp. 2450–2525, 2020.
- [16] 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. 116 753–116 773, 2019.
- [17] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, and L. Hanzo, "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, no. 11, pp. 2450–2525, 2020.
- [18] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active ris vs. passive ris: Which will prevail in 6g?" *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1707–1725, 2023.
- [19] W. Tang, M. Z. Chen, J. Y. Dai, Y. Zeng, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 180–187, 2020.
- [20] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field mimo communications for 6g: Fundamentals, challenges, potentials, and future directions," *IEEE Communications Magazine*, vol. 61, no. 1, pp. 40–46, 2023.
- [21] P. Nepa and A. Buffi, "Near-field-focused microwave antennas: Near-field shaping and implementation," *IEEE Antennas and Propagation Magazine*, vol. 59, no. 3, pp. 42–53, 2017.
- [22] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, and Y. C. Eldar, "6g wireless communications: From far-field beam steering to near-field beam focusing," *IEEE Communications Magazine*, vol. 61, no. 4, pp. 72–77, 2023.
- [23] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of reconfigurable intelligent surfaces (riss): Passive beamforming, modulation, and resource allocation," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4347–4363, 2021.
- [24] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, M. D. Renzo, and F. Chen, "Reconfigurable intelligent surfaces with reflection pattern modulation: Beamforming design and performance analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 741–754, 2021.
- [25] Z. Yang, W. Xu, C. Huang, J. Shi, and M. Shikh-Bahaei, "Beamforming design for multiuser transmission through reconfigurable intelligent surface," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 589–601, 2021.

- [26] J. Y. Dai, J. Zhao, Q. Cheng, and T. J. Cui, "Independent control of harmonic amplitudes and phases via a time-domain digital coding metasurface," *Light: Science & Applications*, vol. 7, no. 1, p. 90, 2018.
- [27] V. S. Asadchy, M. Albooyeh, S. N. Tsvetkova, A. Díaz-Rubio, Y. Ra'di, and S. A. Tretyakov, "Perfect control of reflection and refraction using spatially dispersive metasurfaces," *Physical Review B*, vol. 94, no. 7, p. 075142, 2016.
- [28] F. Wang and A. L. Swindlehurst, "Applications of absorptive reconfigurable intelligent surfaces in interference mitigation and physical layer security," *IEEE Transactions on Wireless Communications*, vol. 23, no. 5, pp. 3918–3931, 2024.
- [29] S. Basharat, M. Khan, M. Iqbal, U. S. Hashmi, S. A. R. Zaidi, and I. Robertson, "Exploring reconfigurable intelligent surfaces for 6g: State-of-the-art and the road ahead," *IET Communications*, vol. 16, no. 13, pp. 1458–1474, 2022.
- [30] T. Bai, C. Pan, Y. Deng, M. El Kashlan, A. Nallanathan, and L. Hanzo, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2666–2682, 2020.
- [31] Z. Chu, P. Xiao, M. Shojafar, D. Mi, J. Mao, and W. Hao, "Intelligent reflecting surface assisted mobile edge computing for internet of things," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 619–623, 2021.
- [32] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 270–278.
- [33] S. Kayraklik, I. Yildirim, I. Hokelek, Y. Gevez, E. Basar, and A. Gorcin, "Indoor measurements for ris-aided communication: Practical phase shift optimization, coverage enhancement, and physical layer security," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 1243–1255, 2024.
- [34] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, O. S. Badarneh, X. Li, and R. Kharel, "Reconfigurable intelligent surface enabled iot networks in generalized fading channels," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [35] Y. Chen, Y. Wang, J. Zhang, P. Zhang, and L. Hanzo, "Reconfigurable intelligent surface (ris)-aided vehicular networks: Their protocols, resource allocation, and performance," *IEEE Vehicular Technology Magazine*, vol. 17, no. 2, pp. 26–36, 2022.
- [36] L. Yang, P. Li, F. Meng, and S. Yu, "Performance analysis of ris-assisted uav communication systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 9078–9082, 2022.
- [37] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410–1414, 2019.
- [38] J. Zuo, Y. Liu, Z. Qin, and N. Al-Dhahir, "Resource allocation in intelligent reflecting surface assisted noma systems," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 7170–7183, 2020.

- [39] C. Pan, H. Ren, K. Wang, M. Elakashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided mimo broadcasting for simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1719–1734, 2020.
- [40] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [41] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser mimo systems exploiting deep reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, 2020.
- [42] M. Samir, M. Elhattab, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3978–3983, 2021.
- [43] K. K. Nguyen, S. R. Khosravirad, D. B. da Costa, L. D. Nguyen, and T. Q. Duong, "Reconfigurable intelligent surface-assisted multi-uav networks: Efficient resource allocation with deep reinforcement learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 358–368, 2022.
- [44] Z. Peng, T. Li, C. Pan, H. Ren, W. Xu, and M. D. Renzo, "Analysis and optimization for ris-aided multi-pair communications relying on statistical csi," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3897–3901, 2021.
- [45] B. Yan, Q. Zhao, J. Zhang, J. A. Zhang, and X. Yao, "Hybrid beamforming for ris-aided communications: Fitness landscape analysis and niching genetic algorithm," *arXiv preprint arXiv:2109.09054*, 2021.
- [46] H. Zhou, M. Erol-Kantarci, Y. Liu, and H. V. Poor, "Heuristic algorithms for ris-assisted wireless networks: Exploring heuristic-aided machine learning," *IEEE Wireless Communications*, vol. 31, no. 4, pp. 106–114, 2024.
- [47] Y.-T. Li, L.-H. Shen, K.-T. Feng, and C.-Y. Chan, "Genetic multi-agent reinforcement learning for multiple double-sided star-ris in full-duplex mimo networks," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 5003–5008.
- [48] M. Polese, M. Dohler, F. Dressler, M. Erol-Kantarci, R. Jana, R. Knopp, and T. Melodia, "Empowering the 6g cellular architecture with open ran," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 2, pp. 245–262, 2024.
- [49] J. Mitola, "The software radio architecture," *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26–38, 1995.
- [50] E. Basar, G. C. Alexandropoulos, Y. Liu, Q. Wu, S. Jin, C. Yuen, O. A. Dobre, and R. Schober, "Reconfigurable intelligent surfaces for 6g: Emerging hardware architectures, applications, and open challenges," *IEEE Vehicular Technology Magazine*, vol. 19, no. 3, pp. 27–47, 2024.

- [51] D. Han, P. Wang, W. Ni, W. Wang, A. Zheng, D. Niyato, and N. Al-Dhahir, "Multi-functional ris integrated sensing and communications for 6g networks," *IEEE Transactions on Wireless Communications*, vol. 24, no. 2, pp. 1146–1161, 2025.
- [52] A. Lambora, K. Gupta, and K. Chopra, "Genetic algorithm- a literature review," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, pp. 380–384.
- [53] Y. Li, "Deep reinforcement learning: An overview," *arXiv preprint arXiv:1701.07274*, 2017.
- [54] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.
- [55] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends® in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.
- [56] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The nonstochastic multiarmed bandit problem," *SIAM journal on computing*, vol. 32, no. 1, pp. 48–77, 2002.

Annexe A - La connexion SPI Daisy chain et RIS

