UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FELIPE AUGUSTO DUTRA BUENO

MASSIVE MIMO: RANDOM ACCESS PROTOCOLS BASED ON STATISTICAL INFERENCE AND MACHINE LEARNING TECHNIQUES

DISSERTAÇÃO

CORNÉLIO PROCÓPIO
2022

FELIPE AUGUSTO DUTRA BUENO

MASSIVE MIMO: RANDOM ACCESS PROTOCOLS BASED ON STATISTICAL INFERENCE AND MACHINE LEARNING TECHNIQUES

MIMO Massivo: Protocolos de Acesso Aleatório Baseados em técnicas de Inferência Estatística e Aprendizagem de Máquina.

Dissertação apresentado(a) como requisito para obtenção do título(grau) de Mestre em Engenharia Elétrica, do Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador(a): Prof. Dr. José Carlos Marinello Filho

CORNÉLIO PROCÓPIO

2022

Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Cornélio Procópio

FELIPE AUGUSTO DUTRA BUENO

MASSIVE MIMO: RANDOM ACCESS PROTOCOLS BASED ON STATISTICAL INFERENCE AND MACHINE LEARNING TECHNIQUES

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Sistemas Eletrônicos Industriais.

Data de aprovação: 20 de Dezembro de 2022

Dr. Jose Carlos Marinello Filho, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Paulo Rogerio Scalassara, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Taufik Abrao, Doutorado - Universidade Estadual de Londrina (Uel)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 20/12/2022.

I dedicate this work to God for granting me health, to my family and friends, for the moments of absence, and to my supervisor for all his patience, dedication, and guidance throughout the work.

ACKNOWLEDGEMENTS

This work was only completed with the help of several people and/or institutions to which I pay tribute. Indeed, these paragraphs will not serve all the people who were part of this important phase of my life. Therefore, I apologize in advance to those who are not present among these words, but they can be sure that they are part of my thoughts and gratitude.

I begin by thanking my advisor José Carlos Marinello Filho, for all his patience and dedication in helping me whenever I needed it, all the knowledge he passed on to me, and without whom the conclusion of this work would not have been possible.

Secondly, I would like to thank everyone who contributed to the preparation of articles related to this work, especially Professor Taufik Abrão, who shared his knowledge with valuable information on how to improve the progress of the works. I also would like to thank Professor Alessandro Goedtel, who encouraged me as his student to do homework always in article format, and his contributions as a co-author of the articles. Finally, I would like to express my gratitude to professor Paulo Scalassara for raising my interest in machine learning methods and statistical inference during his classes.

I also thank all professors and colleagues in the department who helped directly and indirectly in the completion of this work.

Finally, I want to thank my family for their love, encouragement, and total support at all times in my life.

No great work is ever done in a hurry. To develop a great scientific discovery, to print a great picture, to write an immortal poem, to become a minister, or a famous general—to do anything great requires time, patience, and perseverance. These things are done by degrees, "little by little." (W. J. Wilmont Buxton)

RESUMO

BUENO, Felipe Augusto Dutra. **MIMO Massivo: protocolos de acesso aleatório baseados em técnicas de inferência estatística e aprendizagem de máquina**. 2022. 141 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

O número de dispositivos da quinta geração (5G) de redes celulares está crescendo em um ritmo sem precedentes. A tecnologia 5G é caracterizada pela capacidade de prover três serviços essenciais: banda larga móvel aperfeiçoada (eMBB do inglês: *enhanced Mobile Broadband*), comunicação ultra-confiável e com baixa latência (URLLC do inglês: *Ultra-Reliable Low-Latency Communication*), e comunicação máquina-máquina (MTC do inglês: *Machine-To-Machine*) em massa (mMTC do inglês: *massive MTC*). Para suprir essas demandas, diversas tecnologias de suporte para o 5G e seus sucessores (B5G do inglês: *Beyond 5G*) estão sendo desenvolvidas nos últimos anos. Entre essas tecnologias estão as superfícies refletoras inteligentes (IRS do inglês: *Intelligent Reflecting Surface*), acesso múltiplo não ortogonal (NOMA do inglês: *Non-Orthogonal Multiple Access*), ondas milimétricas (mmWave do inglês: *Millimeter-wave*) e múltiplas entradas e múltiplas saídas (MIMO do inglês: *Multiple-Input Multiple-Output*) massivo. Entre todas essas tecnologias, o MIMO massivo é a mais bem sucedida. A tecnologia MIMO massivo é um dos grandes facilitadores da implementação de tecnologia mMTC, cujos dispositivos estarão disponíveis em quantidade massiva e deverão exigir baixo consumo de energia e alta conectividade. Entretanto, visto que os recursos de tempo e frequência fornecidos pelas estações de base (BSs do inglês: *base-station*) são escassos, e a quantidade de dispositivos continua aumentando, é provável que no futuro haverá falta de sinais piloto para atender a todos os dispositivos de uma rede, gerando possíveis gargalos de desempenho no sistema. Para resolver esse problema, alguns protocolos de acesso aleatório vêm sendo desenvolvidos. Este é o caso do protocolo de resolução de colisão de usuário mais forte (SUCRe do inglês: *Strongest-User Collision Resolution*), um protocolo baseado em *concessão de acesso* (GB do inglês: *grant-based*). Outros protocolos *livres de concessão de acesso* (GF do ingês: *grant-free*) mais adequados ao suporte de sistemas mMTC também são propostos. Nete trabalho, três protocolos de acesso aleatório foram propostos. O primeiro é baseado no protocolo SUCRe e tem por objetivo otimizar a etapa de decisão deste protocolo através de um classificador Bayesiano. Este primeiro protocolo apresenta resultados superiores ao SUCRe e de maneira interessante, o efeito de utilizar o classificador Bayesiano somente troca o critério de decisão do protocolo para um mais efetivo. O segundo protocolo é similar ao primeiro, porém utiliza um classificador baseado em uma rede neural perceptron de múltiplas camadas ao invés de um classificar Bayesiano, apresentando resultados ligeiramente superiores. O terceiro protocolo é do tipo GF e utiliza o algoritmo de aprendizado por reforço chamado Q-Learning para guiar os dispositivos em direção a sinais piloto que estejam menos congestionados. Os níveis de congestionamento são obtidos em um cenário de simulação de um sistema MIMO massivo onde os efeitos de propagação realísticos são considerados. O algoritmo se destacou em relação aos métodos tradicionais e às referências de comparação, apresentando melhores resultados nas métricas de taxa de dados de rede, taxa de dados por usuário e latência. Os três protocolos apresentados demonstraram robustez em relação a variação de alguns parâmetros, reafirmando sua eficácia.

Palavras-chave: MIMO Massivo. Acesso Aleatório. Método de Bayes. Q-Learning. B5G.

ABSTRACT

BUENO, Felipe Augusto Dutra. **Massive MIMO: random access protocols based on statistical inference and machine learning techniques**. 2022. 141 p. Dissertation (Master's Degree in Electrical Engineering) – Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

The number of fifth-generation (5G) cellular network devices is growing at an unprecedented rate. 5G technology is characterized by the ability to provide three types of essential services: enhanced Mobile Broadband (eMBB), Ultra-reliable Low Latency Communication (URLLC), and massive MTC (mMTC). These services support many types of applications such as virtual reality, augmented reality, traffic control, Internet of Things (IoT), industrial IoT, and others. To meet these demands, several technologies to support 5G and beyond (B5G) have been developed in recent years. Among these technologies are Intelligent Reflecting Surfaces (IRS), non-orthogonal multiple access (NOMA), millimeter-wave (mmWave), and massive multiple-input multiple-output (MIMO). Of all these technologies, massive MIMO is the most successful. Massive MIMO is a major enabler for the implementation of mMTC technology, whose devices will be available in massive numbers and will require low power consumption and high connectivity. However, since the time and frequency resources provided by Base stations (BSs) are scarce, and the number of devices keeps increasing, it is likely that in the future, there will be a lack of pilot signals to serve all devices in the network, leading to a performance bottleneck in the system. To solve this problem, some random access protocols have been developed. This is the case of the *strongest-user collision resolution* (SUCRe) protocol, a grant-based (GB) protocol that grants access to network resources only to users that are the strongest contender for a particular pilot. Other protocols of the grant-free (GF) type, which is more suitable for supporting mMTC systems, are also proposed. In this work, three random access protocols are proposed. The first one is based on the SUCRe protocol and aims to optimize the decision step of the SUCRe protocol through a Bayesian classifier. This first protocol shows better results than SUCRe, and, interestingly, the effect of using the Bayesian Classifier (BC) only changes the decision criteria of the protocol to a more effective one. The second protocol is similar to the first but uses a classifier based on a Multilayer perceptron (MLP) Neural Network (NN) instead of a Bayesian classifier, presenting slightly better results. The third protocol is a GF Random-Access (RA) procedure and uses a reinforcement learning algorithm called Q-Learning to guide devices toward pilot signals that are less congested. The congestion levels are obtained in a massive MIMO simulation scenario with the realistic propagation effects considered. The algorithm stood out compared to traditional methods and comparison references, showing better results in metrics such as network throughput, per-user throughput, and latency. The three protocols presented also showed robustness in relation to variations of some parameters, reinforcing their efficacy.

Keywords: Massive MIMO. Random Access. Bayes' method. Q-Learning. B5G.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

INITIALISM	
1G	first-generation
2G	second-generation
3G	third-generation
3GPP	3rd generation partnership project
4G	fourth-generation
5G	fifth-generation
5G PPP	Public Private Partnership on 5G
AI	Artificial Intelligence
AMC	Automatic Modulation Classification
ANAA	Average Number of Access Attempts
ANATEL	National Telecommunications Agency of Brazil, from portuguese: *Agência Nacional de Telecomunicações*
AoA	Angle-of-Arrival
AWGN	Additive White Gaussian Noise
B5G	5G and beyond
BC	Bayesian Classifier
BS	Base station
CM	Confusion Matrix
cMBB	crowded Mobile Broadband
CSI	Channel State Information
DL	Downlink
eMBB	enhanced Mobile Broadband
EU	European Unioun
FDD	Frequency-Division Duplex
FDR	false discovery rate
FFAA	Fraction of Failed Access Attempts
FNR	false negative rate
FPC	Fractional Power Control
GB	grant-based
GF	grant-free
HOC	High-Order Cumulant
HTH	Human-to-Human
i.i.d	independent, identically distributed

ICI	Inter-cellular interference
IoT	Internet of Things
IRS	Intelligent Reflecting Surfaces
LDA	Linear Discriminant Analysis
LS	Least Squares
LSTM	Long Short-Term Memory
LTE	Long Term Evolution
MDP	Markov decision process
MF	Matched-Filter
MIMO	multiple-input multiple-output
ML	Machine learning
MLP	Multilayer perceptron
MMSE	Minimum Mean Square Error
mMTC	massive MTC
mmWave	millimeter-wave
MR	Maximum Ratio
MRC	Maximum-Ratio combining
MRT	Maximum-Ratio transmission
MTC	machine type communications
MTM	Machine-to-Machine
MU-MIMO	multi-user MIMO
NB-IoT	narrowband IoT
NN	Neural Network
NOMA	non-orthogonal multiple access
OFDM	Orthogonal Frequency Division Multiplexing
OMA	Orthogonal multiple access
PCA	Principal Component Analysis
PDF	probility density function
PDP	Payload data pilot
PPV	positive predictive value
RA	Random-Access
RACH	Random Access Channel
RAR	Random Access Response
RLe	Reinforcement Learning
RRC	Radio Resource Control
SARSA	State–Action–Reward–State–Action
SBI	Sparse-Bayesian Inference
SI	Statistical Inference
SIC	Successive interference cancellation

SINR	Signal-to-Interference-plus-Noise ratio
SLe	Supervised Learning
SLI	Statistical location information
SNR	Signal-to-Noise Ratio
SOM	Self-Organizing Maps
SORTE	Second-Order Statistic of the Eigenvalues
SUCRe	*strongest-user collision resolution*
SUCRe-BC	SUCRe enhanced by the Bayesian Classifier
SVM	Support Vector Machine
TD	Temporal-Difference
TDD	Time-Division Duplex
TPR	true positive rate
UDN	Ultra Dense Networks
UE	User Equipment
UK	United Kingdom
UL	Uplink
ULe	Unsupervised Learning
URLLC	Ultra-reliable Low Latency Communication
W-CDMA	Wideband code division multiple access
ZF	Zero-Forcing

LIST OF SYMBOLS AND NOTATIONS

SYMBOLS

a	Action taken by an agent in state s.	
a_{t+1}	The next action taken in the next state s_{t+1}.	
$\mathcal{A}_i$	Set of all active User Equipments (UEs) in cell i.	
A_F	Number of failed packet transmission attempts.	
A_T	Number of transmission attempts for transmitting L_k packets.	
$\mathbf{b}_l$	Complex-valued vector channel between the k-th interferer and BS in cell 0.	
$\mathbf{B}$	Complex-valued interference matrix caused in the BS by adjacent cells.	
B_c	Coherence bandwidth.	[Hz]
c	Speed of light.	[m/s]
$c(k)$	Pilot chosen by UE or device k.	
C_j	j-th class or *state of nature* of a given dataset.	
C_{dl}	Point-to-point MIMO downlink channel capacity.	[bits/s]
C^{dl}	MU-MIMO downlink channel capacity.	[bits/s]
C_{ul}	Point-to-point MIMO uplink channel capacity.	[bits/s]
C^{ul}	MU-MIMO uplink channel capacity.	[bits/s]
$C_L(k)$	Number of contenders for pilot $c(k)$.	
$\mathbf{d}_l$	Complex-valued random data sequence transmitted by the interferer.	
$d\mathbf{x}_d$	d-space volume element.	
$\mathbf{d}_k$	Complex-valued payload data vector.	
$\widehat{\mathbf{d}}_k$	Estimator for the complex-valued payload data vector.	
$dist_k$	Distance between the k-th device and its BS.	
E_p	The length of one episode.	
$\mathbf{F}$	Complex-valued linear precoding matrix.	
$\mathbf{f}_i$	i-th column vector of the linear precoding matrix $\mathbf{F}$.	
$\mathbf{g}$	M-dimensional channel response.	
$\mathbf{g}(\mathbf{t})$	Generic wireless channel as function of time t.	
$\mathbf{G_{dl}}$	Complex-valued downlink wireless channel matrix.	
$\mathbf{G_{ul}}$	Complex-valued uplink wireless channel matrix.	
$\mathbf{g}_{\mathrm{dl}_k}$	k-th row vector of the matrix $\mathbf{G_{dl}}$.	
g_k^m	(m,k)-th element of $\mathbf{G_{ul}}$.	
$\hat{g}_k^m$	Estimator for g_k^m.	
$\mathbf{g_{ul}}_j$	j-th column vector of $\mathbf{G_{ul}}$.	

$\mathbf{g}_{t,k}$	Channel between the member $k \in \mathcal{S}_t^{\text{interf}}$ and the BS in cell 0.	
$grad_j^{(l)}$	Local gradient with respect to the j-th neuron of the l-th layer.	
$\mathbf{h}_k$	Complex-valued vector channel between the k-th UE or device and its BS.	
$\mathcal{I}_k$	Condition for the UE k do not retransmit its pilot.	
$\mathcal{K}_0$	Set of inactive UEs in cell 0.	
K_0	Number of inactive UEs in cell 0.	
K_a	Number of active devices in the cell.	
$\widehat{K}_a$	Estimator for the number of active devices in the cell.	
K_T	Total number of devices in a cell.	
K_{ici}	Number of interfering devices in each neighboring cell.	
L	Number of packets.	
L_k	Number of packets that device k have to send.	
L_I	Number of neurons in the input layer.	
L_H	Number of neurons in the hidden layer.	
L_O	Number of neurons in the output layer.	
M	Number of BS antennas.	
$\mathbf{N}_d$	Data receiver noise matrix, with entries independent, identically distributed (i.i.d), following a complex normal distribution, $\mathcal{CN}(0,1)$.	
$\mathbf{N}_p$	Pilot receiver noise matrix, with entries i.i.d, following a complex normal distribution, $\mathcal{CN}(0,1)$.	
$\mathbf{N}'_p$	Correlation matrix of $\mathbf{N}_p$ with the pilot sequences.	
$[\mathbf{N}'_p]_{mk}$	(m,k)-th element of $\mathbf{N}'_p$.	
$\mathbf{n_k}$	Perceived noise at the k-th terminal antenna.	
$\mathbf{n_{ul}}$	Complex-valued uplink noise plus interference vector.	
$\mathbf{n_{dl}}$	Complex-valued downlink noise plus interference vector.	
$\mathbf{n}_t$	Effective noise with distribution $\mathcal{CN}(0, \sigma^2\mathbf{I}_M)$.	
$\mathbf{o}_k$	Desired output vector for input sample k.	
$\hat{\mathbf{o}}_k$	Output vector for input k.	
$o_j^{(l)}$	The value of the j-th output provided by the network neuron of the $(l+1)$-th layer.	
P	Number of input samples.	
P_L	Shadow fading at a reference distance.	
P_a	Probability of the UEs or device activation.	
P_r	Probability of retransmitting a pilot when a transmission fail occurs.	
$P_c(k)$	Collaborative penalty funcion.	
q	Downlink transmit power.	[dBm]
r	Immediate reward for taking action a in state s.	

R	Hexagonal cell radius.	
$\mathcal{R}_k$	Condition for the UE k retransmit its pilot.	
$\mathcal{R}$	Set of rewards.	
$R_F(k,l)$	Reward function of a user k choosing a pilot l.	
s	Current state of an agent.	
s_{t+1}	The next state of an agent after taking action a.	
$\mathcal{S}_t$	Set of indices of UEs that transmit pilot t.	
$\mathcal{S}_t^{\text{interf}}$	Set of interfering UEs in neighboring cells that reuses pilot t.	
T	Number of terminal antennas.	
T_c	Coherence time.	[s]
T_s	Time slot.	
$\mathbf{U}$	Weight matrix.	
$u_{ji}^{(l)}(t)$	The synaptic weight between the j-th neuron of the output layer $(l+1)$ and the i-th neuron of layer l.	
$u_{ji}^{(l)}(t+1)$	Updated version of $u_{ji}^{(l)}(t)$.	
$\mathcal{U}_i$	Set of all UEs in cell i.	
$\mathbf{v}$	Beamforming vector.	
$\mathbf{V}$	Precoded Downlink (DL) pilot signal matrix .	
$\mathbf{W}$	Complex-valued detector matrix.	
$\mathbf{w}_i$	i-th row vector of the matrix $\mathbf{W}$.	
$\mathbf{x}$	Complex-valued signal or data vector.	
$\mathbf{x}_d^R$	d-Dimensional random vector.	
$\mathbf{x}_d$	d-Dimensional signal or data vector.	
$\mathbf{x_{ul}}$	Complex-valued uplink signal vector.	
$\mathbf{x_{dl}}$	Complex-valued downlink signal vector.	
$\hat{\mathrm{x}}_{\mathrm{ul}}^i$	i-th symbol received by the BS.	
$\mathbf{y_{ul}}$	Complex-valued uplink signal vector received by the BS.	
$\mathbf{y_{dl}}$	Complex-valued downlink signal vector received by the T terminal antennas.	
$\hat{\mathrm{y}}_{\mathrm{dl}_k}$	Signal received by the k-th terminal antenna.	
$\mathbf{X}_p$	Complex-valued matrix of pilot signal sent by the terminals.	
$\hat{\mathbf{x}}_{\mathbf{ul}}$	Uplink signal vector estimation.	
$\hat{\mathrm{x}}_{\mathrm{ul}}^i$	i-th received symbol at the BS.	
$\mathbf{Y}$	Complex-valued received signal matrix.	
$\mathbf{Y_d}$	Complex-valued received data signal matrix.	
$\mathbf{Y}_p$	Complex-valued matrix of pilot signal received by the BS.	
$\mathbf{y}_t$	Correlation of the signal $\mathbf{Y}_p$ received by the BS with pilot ψ_t.	
$\mathbf{Y}'_p$	Correlation matrix of $\mathbf{Y}_p$ with the pilot sequences.	
$[\mathbf{Y}'_p]_{mk}$	(m,k)-th element of $\mathbf{Y}'_p$.	
$\mathbf{z}_k$	$\tau_p \times 1$ complex-valued signal received by the k-th UE or device during downlink stage of the SUCRe protocol.	

z_k	Correlation of $\mathbf{z}_z$ with the pilot ψ_t chosen by the UE or device k.	
α_t	Sum of the signal strengths and interference received by the BS for each pilot t.	
$\hat{\alpha}_{t,k}$	Estimator for the sum of the signal strengths and interference received by the BS for each pilot t estimated by k-th UE or device.	
β_k	Large-scale fading coefficient of the k-th uplink channel.	
γ	Q-Learning discount factor.	
γ_k^{ul}	Uplink signal to interference plus noise ratio of UE or device k.	
δ	Number of standard deviations.	
ϵ_k	Bias parameter of the SUCRe protocol associated with the decision conditions of UE k.	
$\boldsymbol{\eta}_k^{\mathrm{T}}$	Noise signal with distribution $\mathcal{CN}(0,\sigma^2\mathbf{I}_{\tau_p})$, received during downlink stage of the SUCRe protocol.	
η_k	Effective receiver noise during downlink stage of the SUCRe protocol, with distribution $\mathcal{CN}(0,\sigma^2)$.	
$\boldsymbol{\theta}$	Power allocation vector.	
ϑ	Path loss exponent.	
κ	Learning rate.	
λ	Wavelength.	$[\mathrm{m}]$
μ	Mean value.	
ν	Devices's velocity.	$[\mathrm{m/s}]$
$\boldsymbol{\nu}_k^{\mathrm{T}}$	Complex-valued intercell-interference vector term perceived by the k-th UE or device during downlink stage of the SUCRe protocol.	
ξ	Training precision.	
π	Greedy policy.	
ρ	Uplink channel SNR	$[\mathrm{dBm}]$
ρ_{ul}	Uplink transmit power.	$[\mathrm{dBm}]$
ρ_i	Power of transmission of a pilot signal by an UE or device i.	
$\rho_{t,k}$	Transmit power of the member $k \in \mathcal{S}_t^{\text{interf}}$.	
ρ_{snr}	SNR at the receiver side.	
σ	Standard deviation.	
σ^2	Variance.	
$\sigma_{\mathrm{s}f}$	Shadow fading standard deviation.	
τ_c	Number of coherence blocks.	$[\mathrm{samples}]$
τ_p	Number coherence blocks dedicated to pilots.	$[\mathrm{samples}]$
τ_{dl}	Number coherence blocks dedicated to downlink data.	$[\mathrm{samples}]$
$\tau_{dl,p}$	Number coherence blocks dedicated to downlink pilots.	$[\mathrm{samples}]$
τ_{ul}	Number coherence blocks dedicated to uplink data.	$[\mathrm{samples}]$
$\tau_{ul,p}$	Number coherence blocks dedicated to uplink pilots.	$[\mathrm{samples}]$
ϕ_k	Proportion of the signal strength of the k-th UE among the pilot contenders.	

Φ	Scalar gain.
φ	Shadow fading.
$\boldsymbol{\psi_i}$	i-th pilot sequence.
$\boldsymbol{\Psi}$	Matrix of pilot sequences (as its columns).
ω	Interference.
Ω_k	The class state of the UE k.
$\widehat{\Omega_k}$	Estimator of Ω_k.

NOTATIONS

$\mathbb{C}$	Space of complex-valued numbers.
$\mathbb{R}$	Space of real-valued numbers.
$\mathbf{X}^{\mathrm{T}}$	The transpose of a matrix $\mathbf{X}$.
$\mathbf{X}^{\mathrm{H}}$	The conjugate-transpose of a matrix $\mathbf{X}$.
$\mathbf{X}^{*}$	The conjugate of a matrix $\mathbf{X}$.
$\mathbf{I}_M$	$M \times M$ identity matrix.
$\mathbf{D}_{\mathbf{a}}$	Diagonal matrix with elements of vector $\mathbf{a}$ as its diagonal elements.
$a \propto b$	a is proportional to b .
$a > b$	a is greater than b .
$a \gg b$	a is much greater than b .
$a \geq b$	a is greater than or equal to b .
$a < b$	a is less than b.
$a \leq b$	a is less than or equal to b.
$\mathcal{A} \subseteq \mathcal{B}$	$\mathcal{A}$ is a subset of $\mathcal{B}$.
$f(x)$	Function$f : X \to Y$ with $x \in X$.
$f'(x)$	Derivative of function $f(x)$.
$\lambda(\cdot)$	Loss function.

$\max(f(x))$	The maximum value of the function $f(x)$.
$\max(a,b)$	Return a if $a \geq b$ or b if $b \geq a$.
$\min(f(x))$	The minimum value of the function $f(x)$.
$\min(a,b)$	Return a if $a \leq b$ or b if $b \leq a$.
$sig(\cdot)$	Sigmoid function.
$\mathcal{B}(N_t,P_s)$	Binomial distribution, where N_t is the number of trials and P_s is the success probability of each trial.
$Q(s,a)$	Cummulative reward function of an agent in the state s taking an action a.
$Q(s,a)$	Cummulative reward function of an agent in the state s taking an action a.
$\mathbb{E}[\cdot]$	The expectation with respect to a random variable.
$\mathbb{V}[\cdot]$	The variance with respect to a random variable.
$\bar{x}_k$	Normalization of scalar x_k.

$\bar{\mathbf{x}}_k$	Normalization of vector $\mathbf{x}_k$.
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution, with mean μ and standard deviation σ.
$\mathcal{CN}(\mu, \sigma^2)$	Circularly-symmetric complex Gaussian distribution. Such that, if $x \sim \mathcal{CN}(\mu, \sigma^2)$ implies that $x \sim \mathcal{N}(\mu_R, \frac{\sigma^2}{2}) + i \cdot \mathcal{N}(\mu_I, \frac{\sigma^2}{2})$, with $\mu = \mu_R + i \cdot \mu_I$.
$\|a\|$	The magnitude $a \in \mathbb{C}$ or cardinality of a if a is a set.
$\|\cdot\|$	The 2-norm.
$\Gamma(\cdot)$	The Gamma function.
$\log_b(\cdot)$	Log base b function.
$\ln(\cdot)$	Natural log function.
$\det(\mathbf{A})$	The determinant of matrix $\mathbf{A}$.
$P(A)$	Probability of event A occurring.
$P(A\|B)$	Conditional probability. The probability of event A occurring given that event B has occurred.
$P(A,B)$	Joint probability of A and B.
$R(a\|b)$	Conditional risk of a given b.
$p_x(x)$	Probility density function (PDF) of a continuous random variable x.
$p_x(x\|y)$	Conditional PDF of a continuous random variable x given y.
$\{\mathcal{F} \backslash \mathcal{G}\}$	$\{f \in \mathcal{F} \| f \notin \mathcal{G}\}$

CONTENTS

1 INTRODUCTION

Compared to previous generations, the number of fifth-generation (5G) mobile communication devices has been increasing rapidly all over the world (ERICSSON, 2022). By the end of 2021, more than 20 service providers had launched public 5G standalone networks, and this number is expected to double in 2022 (ERICSSON, 2022). The number of 5G subscribers is estimated to exceed one billion devices in 2022 and reach 4.4 billion by the end of 2027 (ERICSSON, 2022). It is anticipated that by 2023, there will be 13.1 billion mobile devices connected, with 1.4 billion of those ready for 5G networks (CISCO, 2020). According to Mahmood *et al.* (2020, p. 3), the Internet of Things (IoT) connected device count is projected to increase three-fold, from approximately 11 billion in 2019 to 30 billion by 2030. The fast expansion of machine type communications (MTC) devices is being driven by new market trends, including Smart Cities, Industry 5.0, and Artificial Intelligence (AI) aided Autonomous mobility (MAHMOOD *et al.*, 2020).

The growing demand for applications such as high-definition video streaming, edge computing, virtual reality, and MTC systems requires 5G and beyond (B5G) technologies to offer services like enhanced Mobile Broadband (eMBB), Ultra-reliable Low Latency Communication (URLLC) and massive MTC (mMTC). To meet these requirements, disruptive technologies like millimeter-wave (mmWave) spectrum, Intelligent Reflecting Surfaces (IRS), and massive multiple-input multiple-output (MIMO) have been developed, as discussed in (AL-FALAHY; ALANI, 2017; WU; ZHANG, 2019). Among the B5G technologies, massive MIMO is widely recognized as a key component (BJÖRNSON *et al.*, 2019). This technology benefits from the channel hardening effect, which occurs when uncorrelated noise and fast fading disappear due to an increase in the number of transmitting antennas to infinity (NGO; LARSSON, 2017). Theoretical studies have shown that such systems can significantly improve wireless network connection reliability, transmission energy, and spectral efficiency (MARZETTA, 2010; LARSSON *et al.*, 2014). Empirical studies using measured data support these predictions, demonstrating that similar results can be achieved in physical propagation channels (GAO *et al.*, 2015).

B5G systems must also be ready to provide crowded Mobile Broadband (cMBB) services, which becomes necessary when the Base station (BS) cannot serve all devices within the cell, at the same time, due to the lack of orthogonal training signals (pilots) (AFSHAR; VAKILI, 2021; BJÖRNSON *et al.*, 2017a). This scenario gives rise to performance issues,

especially *pilot collision*, which is a problem that occurs when two or more devices choose the same pilot to access the BS resources (BJÖRNSON *et al.*, 2017a). Establishing a Random-Access (RA) policy has been mandatory to address this issue. A RA protocol can be established in two principal forms, either as grant-free (GF) or a grant-based (GB) handshake procedure. In a GB handshake procedure, the device must be granted permission to access the BS and then be allowed to send payload data through an exclusive communication channel. Thus, GB protocols are preferable for eMBB systems since their latency requirements are moderate (NOMEIR *et al.*, 2021). On the other hand, in a GF process, the device will send a pilot only for channel estimation, followed by its payload data. If the BS successfully decodes the payload data, then it informs the device that the transmission succeeded. This approach is better for URLLC systems and devices that usually send small amounts of data, like sensors, as the signaling overhead is lower (FEHRENBACH *et al.*, 2018).

In Björnson *et al.* (2017a), the *strongest-user collision resolution* (SUCRe) protocol is introduced. This protocol is a GB, 4-steps RA procedure whose idea is to grant access, to the BS resources, only to the strongest pilot contender each time. The SUCRe protocol accomplishes this by exploring the massive MIMO channel properties to estimate the sum of all devices' signal strength. In numerical Monte-Carlo simulations, the SUCRe protocol has been shown to be very successful, solving about 90% of all collisions. Other works such as (HAN *et al.*, 2017a; MARINELLO; ABRÃO, 2019; HAN *et al.*, 2017b; MARINELLO *et al.*, 2020) present proposals for optimization of the SUCRe protocol showing promising results. In (BAI *et al.*, 2021), a GF approach that exploits the massive MIMO hardening effect and favorable propagation characteristics are proposed to decode signals even when a pilot collision between 2 or 3 pilot signals occurs. Other GF RA protocols are proposed in (HAN *et al.*, 2020; DING *et al.*, 2019; DING; CHOI, 2020; YU *et al.*, 2020) with promising results in terms of reduced signaling overhead for the devices and the BS at the cost of potential collisions.

Machine learning (ML) approaches are another promising area of research for B5G and other communication technologies. Because of their extremely versatile nature, ML techniques can be applied to a wide range of problems. ML algorithms are classified into three main types: Supervised Learning (SLe), Unsupervised Learning (ULe), and Reinforcement Learning (RLe). In (ZHOU *et al.*, 2018), a SLe approach, specifically, the supervised version of *deep learning*, is used for traffic control in 5G Ultra Dense Networks (UDN), outperforming conventional methods. Other studies such as (HUANG *et al.*, 2018; WEN *et al.*, 2018; CHUN *et al.*, 2019; HUANG *et*

al., 2019) apply RLe-based algorithms in the context of massive MIMO systems, with remarkable results in a variety of areas, including channel estimation, resource assignment, and precoding. RLe based techniques are also an important research topic since they allow systems to learn from the environment without the need for a data training set. An important RLe algorithm called *Q-Learning* has gained attention for its simplicity and has been successfully used in works such (SHARMA; WANG, 2019; BELLO *et al.*, 2014) for reducing pilot collisions in GF RA protocols. In (SANTOS *et al.*, 2022), the *Q-Learning* algorithm is applied in a multi-cellular massive MIMO system for power and pilot allocation showing promising performance and complexity trade-off results in more than one use case scenario. Another interesting SLe approach is the Bayesian method, also known as Bayesian Classifier (BC). The Bayesian method is a classical Statistical Inference (SI) tool that compares the weighted PDFs of various classes and their associated costs and then chooses the class with the highest probability of being true (HASTIE *et al.*, 2009). Many works have successfully applied the Bayesian method to wireless network problems, including (OROZA *et al.*, 2021; SEO *et al.*, 2021; TAO *et al.*, 2021) and others.

1.1 OBJECTIVES

The primary objective of this work is to design and compare GF and GB RA protocols for resolving pilot collisions in massive MIMO systems. To accomplish this goal, the following specific objectives have been set:

- Simulate massive MIMO systems in numerical simulation software, considering realistic propagation effects based on valid mathematical models.
- Minimize the occurrence of *pilot collision* by applying GF and GB protocols to massive MIMO systems.
- Investigate the use of different ML and SI approaches.
- Examine how different approaches affects performances metrics such as *throughput*, *per-user throughput*, *latency*, *average number of access attempts*, and *fraction of failed access attempts* and propose ways to optimize these metrics.

1.2 JUSTIFICATION

The global demand for innovative B5G technologies continues to grow as the fields of MTC and machine learning continues to advance and expand for both industrial and academic applications (ERICSSON, 2022). Among these technologies, massive MIMO has already proven successful in commercial applications and is a critical component of B5G (BJÖRNSON *et al.*, 2019). In Brazil, the use of 5G radio frequencies was officially granted to companies through a government auction in December 2021, with the requirement for nationwide coverage by 2029 (CIVIL HOUSE, 2021). To meet this demand, the National Telecommunications Agency of Brazil, from portuguese: *Agência Nacional de Telecomunicações* (ANATEL) estimates that operators will invest R$ 51 billion over 20 years (Ministry of Communication, 2021). Therefore, there is a local and global necessity for research, such as this work, that proposes optimization techniques for B5G networks.

1.2.1 Related Works

To support this work, three articles were prepared for submission to peer-reviewed journals. Each article proposes a RA protocol as described below:

- In the first article (Appendix A), the proposed protocol is a GB RA approach. In this protocol, a BC is developed to estimate, in a decentralized way, the likelihood of a device being the strongest contender for a particular pilot signal. The proposed BC is trained with data collected through Monte-Carlo realizations of the SUCRe protocol (BJÖRNSON *et al.*, 2017a) and then applied in the third step of the SUCRe protocol. The proposed method differs from the original SUCRe protocol as the number of false-negative cases is highly diminished. The first article, titled *"A Random Access Protocol for Crowded Massive MIMO Systems Based on a Bayesian Classifier"* is now published in the IEEE Wireless Communications Letters journal (BUENO *et al.*, 2022).

- In the second article (Appendix B), the proposed protocol implements a Q-Learning algorithm designed to minimize the number of pilot collisions in a GF RA scenario mainly developed for mMTC systems. The implemented protocol is based on the collaborative multi-agent algorithm proposed in (SHARMA; WANG, 2019). However, the implemented protocol is applied in a scenario where realistic propagation effects such as multipath

fading, shadowing, path loss, thermal noise, and Inter-cellular interference (ICI) are considered. The second article was submitted to the journal "Computer Networks" (JCR = 5.493), published by Elsevier, on November 17, 2022, and is currently under revision state R0.

- In the third article (Appendix C), a Neural Network (NN) GB RA protocol is proposed. This protocol is similar to the BC classifier presented in the first article (BUENO *et al.*, 2022). However, a Multilayer perceptron (MLP) NN is employed instead of a BC classifier for predicting the condition of a device being the strongest contender. Similarly to the protocol presented in (BUENO *et al.*, 2022), this protocol is also trained with data collected through Monte-Carlo realizations of the SUCRe protocol (BJÖRNSON *et al.*, 2017a). It differs from the BC as it presents slightly better results. The third article is currently being prepared for the first submission.

2 LITERATURE REVIEW

This chapter reviews the main topics related to massive MIMO systems, random access protocols, and the application of machine learning and statistical inference tools in these protocols.

2.1 5G AND B5G SYSTEMS

The reuse of frequency spectrum at spatially separated locations, made possible by the power fall over distance, is the most significant advantage of cellular networks (GOLDSMITH, 2005). As a result, a BS can be placed in the center of a cell to serve a given region. The cells of first-generation (1G) mobile networks were giant, capable of covering extensive areas (GOLDSMITH, 2005). In 1G networks, the BSs transmitted analog signals at extremely high power levels, making BS installation prohibitively expensive. Given that digital systems were cheaper and easier to implement than analog systems, the second-generation (2G) introduced them to the cellular networks (GOLDSMITH, 2005). The third-generation (3G) introduced Wideband code division multiple access (W-CDMA) and were able to provide much faster data rates than the previous generations (GOLDSMITH, 2005). The fourth-generation (4G) of mobile networks established the Long Term Evolution (LTE) standard, promising to improve the provided data rates even further. To achieve this goal, 4G systems implemented MIMO and Orthogonal Frequency Division Multiplexing (OFDM) technologies. One of the first steps towards making 5G the new mobile communication standard was taken in 2012 by the European Unioun (EU) that officially launched the mobile and wireless communications enables for the 2020 information society (METIS) project to conduct research on 5G mobile communication networks (WANG; MA, 2019). In 2013, the European Commission launched the Public Private Partnership on 5G (5G PPP) to conduct research in 5G (5G PPP, 2013). Countries such as United Kingdom (UK), South Korea, and China have also established their own 5G research projects. Nowadays, 5G is already considered the new global mobile communication standard (WANG; MA, 2019).

There are three main usage scenarios envisioned for 5G systems: eMBB, URLLC, and mMTC. To provide these services, 5G systems must meet a whole new set of requirements. The main user experience requirements of 5G systems are presented in (NGMN, 2015) and summa-

rized in Table 1. B5G system research aims to improve the previously mentioned services even further while also considering future challenges by developing new and disruptive technologies such as the IRS, mmWaves, and others. Currently, the most successful B5G technology is the massive MIMO which has not only been the subject of many research works but has also found a lot of industrial applications, including in mMTC use cases (BJÖRNSON *et al.*, 2019).

Table 1 – User Experience Requirements

Use case category	User Experienced Data Rate	E2E Latency	Mobility
Broadband access in dense areas	DL: 300 Mbps UL: 50 Mbps	10 ms	On demand, 0-100 km/h
Indoor ultra-high broadband access	DL: 1 Gbps UL: 500 Mbps	10 ms	Pedestrian
Broadband access in a crowd	DL: 25 Mbps UL: 50 Mbp	10 ms	Pedestrian
50+ Mbps everywhere	DL: 50 Mbps UL: 25 Mbps	10 ms	0-120 km/h
Ultra-low-cost broadband access for low ARPU areas	DL: 10 Mbps UL: 10 Mbps	50 ms	on demand: 0- 50 km/h
Mobile broadband in vehicles (cars, trains)	DL: 50 Mbps UL: 25 Mbps	10 ms	On demand, up to 500 km/h
Airplanes connectivity	DL: 15 Mbps per user UL: 7.5 Mbps per user	10 ms	Up to 1000 km/h
Massive lowcost , long-range and low-power MTC	Low (typically 1-100 kbps)	Seconds to hours	On demand: 0- 500 km/h
Broadband MTC	**See the requirements for the Broadband access in dense areas and 50+Mbps in everywhere categories**		
Ultra-low latency	DL: 50 Mbps UL: 25 Mbps	<1 ms	Pedestrian
Resilience and traffic surge	DL: 0.1-1 Mbps UL: 0.1-1 Mbps	Regular communication: not critical	0-120 km/h
Ultra-high reliability & Ultra-low latency	DL: From 50 kbps to 10 Mbps UL: From a few bps to 10 Mbps	1 ms	On demand: 0- 500 km/h
Ultra-high availability & reliability	DL: 10 Mbps UL: 10 Mbps	10 ms	On demand, 0-500 km/h
Broadcast like services	DL: Up to 200 Mbps UL: Modest (e.g. 500 kbps)	<100 ms	On demand: 0-500 km/h

Source: (NGMN, 2015)

2.2 MASSIVE MACHINE TYPE COMMUNICATIONS

Previous generations of cellular networks were able to provide Human-to-Human (HTH) type of communications, but the support to Machine-to-Machine (MTM) type was limited to a small number of devices in each cell (WANG; MA, 2019). B5G technologies shift this paradigm by providing services that are mMTC ready, making possible a huge variety of applications, including smart agriculture, traffic control, wireless factory automation, internet of drones, and many others (MAHMOOD *et al.*, 2020; WANG; MA, 2019). To do so, B5G must integrate the already provided HTH services with a massive number of low-complexity, low-power machine-type devices (SENEL; LARSSON, 2018). The main obstacles for providing mMTC services are achieving massive connectivity, which is suggested by the 3rd generation partnership project (3GPP) as 10^6 $[\mathrm{devices}/\mathrm{km}^2]$ (WANG; MA, 2019; 3GPP, 2016), and allowing the implementation of low-power consuming devices. Given that a mMTC grid will supposedly be constituted by numerous devices, it is not practical to allow a high energy consumption because the batteries have to be changed frequently, leading to poor system efficiency (WANG; MA, 2019). Massive MIMO have been shown to be a key technology for the integration of HTH, and MTM type communications (SENEL; LARSSON, 2018) and the support of mMTC requirements.

2.3 MASSIVE MIMO

The concept of massive MIMO can be regarded as an extension of the MIMO concept introduced by Marconi more than a century ago (ALEXANDERSON, 1919). The benefits of the massive MIMO technology were first demonstrated in (MARZETTA, 2010), where it was shown how increasing the number of transmitting antennas leads to the *channel hardening* effect, which occurs when the number of transmitting antennas M grows to infinity. As a result, the effects of uncorrelated noise and fast fading disappear, leaving only inter-cellular interference caused by pilot contamination. According to Marzetta *et al.* (2016, p. 13), the main importance of the channel hardening effect is that the effective channel between each terminal and the BS can be reduced to a deterministic number. Considering a terminal with M-dimensional channel response $\mathbf{g}$ and a beamforming vector $\mathbf{v}$, the terminal sees a scalar channel with gain $\mathbf{v}^{\mathrm{T}}\mathbf{g}$. When M is large, by virtue of the law of large numbers, $\mathbf{v}^{\mathrm{T}}\mathbf{g}$ is close to its expected value, $\mathbb{E}(\mathbf{v}^{\mathrm{T}}\mathbf{g})$ (a deterministic number). This means that the resulting effective channel between each terminal and

the base station is a scalar channel with known, frequency-independent gain and additive noise. In practice, the channel hardening effect reduces the channel between the terminal k and BS to a deterministic value dependent only on the large-scale fading coefficient, β_k. This is a positive real number that embodies range-dependent path loss and shadow fading. It is virtually independent of frequency and is strongly correlated over many wavelengths of space (MARZETTA *et al.*, 2016). Given the slow-changing nature of the large-scale fading coefficient, it is assumed to be known by the terminal.

The channel hardening effect in massive MIMO systems provides the following benefits: **a)** The scalar channel experienced by each terminal behaves like an Additive White Gaussian Noise (AWGN) channel, enabling the use of standard coding and modulation techniques designed for AWGN channels; **b)** Simple resource allocation and power control schemes can be implemented; **c)** The channel hardening allows for the estimation of average channel gain, which is dependent on β_k, at terminals, without the need for Downlink (DL) pilot transmissions in most cases. The value of β_k is assumed to be known by terminal k, which estimates it based on information periodically received from the BS (MARZETTA *et al.*, 2016).

Before diving deep into the specific details of massive MIMO technology, this section introduces the notion of point-to-point MIMO and multi-user MIMO (MU-MIMO).

2.3.1 Point-to-Point MIMO

The most basic form of MIMO is the point-to-point MIMO. In a point-to-point MIMO system, the BS is equipped with an array of M antennas to serve a single terminal equipped with an array of T antennas. In such a scheme, different terminal devices must be served through a different combination of time and frequency resources. Furthermore, the BS must acquire Channel State Information (CSI) during the Uplink (UL), and the terminal during the downlink (MARZETTA, 2015). A representative illustration of the point-to-point MIMO is depicted in Figure 1.

During the UL, the BS receives a signal $\mathbf{y_{ul}} \in \mathbb{C}^{M\times 1}$ through a wireless propagation channel $\mathbf{G_{ul}} \in \mathbb{C}^{M\times T}$, expressed as:

$$\mathbf{y_{ul}} = \sqrt{\rho_{ul}}\mathbf{G_{ul}}\mathbf{x_{ul}} + \mathbf{n_{ul}}, \tag{1}$$

where ρ_{ul} is the transmit power, $\mathbf{x_{ul}} \in \mathbb{C}^{T\times 1}$ is the uplink signal vector and $\mathbf{n_{ul}} \in \mathbb{C}^{M\times 1}$ is the

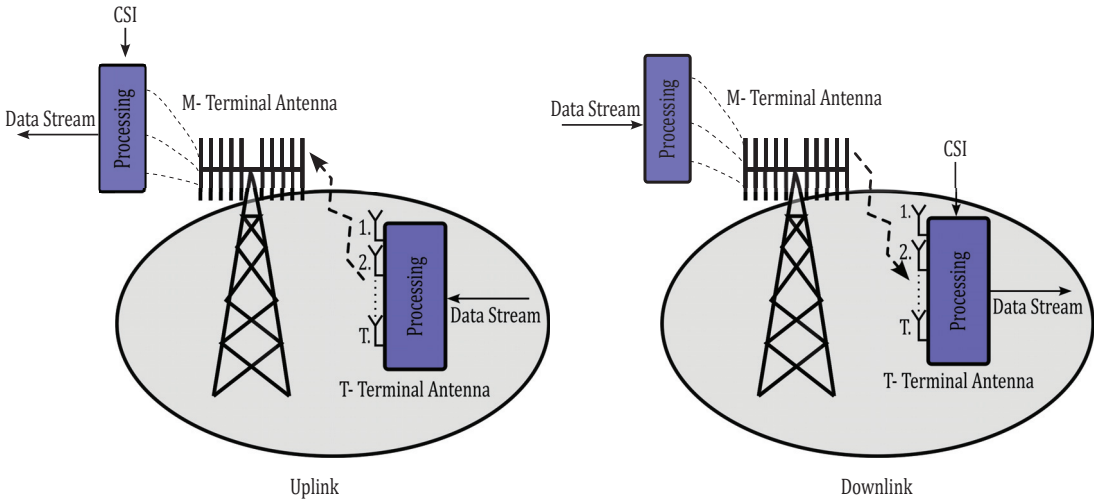


Figure 1 – Illustrative representation of the point-to-point MIMO system

uplink additive noise plus interference.

During the UL the BS decodes the received signal $\mathbf{y_{ul}}$ by employing a detection technique. If a Maximum-Ratio combining (MRC) detector is employed, then the received signal is multiplied by the conjugate-transpose of the UL channel matrix, $\mathbf{G_{ul}}$. If a Zero-Forcing (ZF) detector is employed, then $\mathbf{y_{ul}}$ is multiplied by the UL channel matrix pseudo-inverse. We can generically write any linear detector matrix as $\mathbf{W} \in \mathbb{C}^{T \times M}$ and expand the received symbols to fully characterize them in terms of the desired signal, interference, and noise,

$$\hat{\mathbf{x}}_{\mathbf{ul}} = \mathbf{W}\mathbf{y}_{\mathbf{ul}}, \tag{2}$$

where $\hat{\mathbf{x}}_{\mathbf{ul}}$ represents the UL signal vector estimated by the BS. Expanding the i_{th} received symbol $\hat{\mathrm{x}}_{\mathrm{ul}}^{i}$,

$$\hat{\mathrm{x}}_{\mathrm{ul}}^{i} = \mathbf{w}_i\mathbf{y}_{\mathbf{ul}} = \mathbf{w}_i \cdot (\sqrt{\rho_{ul}}\mathbf{G}_{\mathbf{ul}}\mathbf{x}_{\mathbf{ul}} + \mathbf{n}_{\mathbf{ul}}) = \mathbf{w}_i \sum_{j=1}^{T} \sqrt{\rho_{ul}}\mathrm{x}_j\mathbf{g}_{\mathbf{ul}j} + \mathbf{w}_i \cdot \mathbf{n}_{\mathbf{ul}} =$$

$$\underbrace{\sqrt{\rho_{ul}}\mathbf{w}_i\mathbf{g}_{\mathbf{ul}i}\mathrm{x}_i}_{\text{desired signal}} + \underbrace{\sum_{j=1, j\neq i}^{T} \sqrt{\rho_{ul}}\mathbf{w}_i\mathrm{x}_j\mathbf{g}_{\mathbf{ul}j}}_{\text{interference}} + \underbrace{\mathbf{w}_i \cdot \mathbf{n}_{\mathbf{ul}}}_{\text{noise}}, \tag{3}$$

where $\mathbf{w}_i$ represents the i_{th} row vector of the matrix $\mathbf{W}$, $\mathbf{g}_{\mathbf{ul}j}$ indicates the j_{th} column vector of the UL channel matrix $\mathbf{G_{ul}}$ and x_j indicates the j_{th} symbol sent by the transmitter.

The T terminal antennas receives the DL signal $\mathbf{y_{dl}} \in \mathbb{C}^{T \times 1}$ in a similar fashion as:

$$\mathbf{y}_{\mathbf{dl}} = \sqrt{q}\mathbf{G}_{\mathbf{dl}}\mathbf{x}_{\mathbf{dl}} + \mathbf{n}_{\mathbf{dl}}, \tag{4}$$

where $\mathbf{q}$ is the DL transmit power, $\mathbf{x_{dl}} \in \mathbb{C}^{M \times 1}$ is the DL signal vector, $\mathbf{n_{dl}} \in \mathbb{C}^{T \times 1}$ is the DL noise plus interference and $\mathbf{G_{dl}} \in \mathbb{C}^{T \times M}$ is the DL channel matrix.

During the DL the BS usually send a precodified signal vector $\mathbf{x_{dl}}$, with $\mathbf{x_{dl}} = \mathbf{F} \cdot \mathbf{x}$, in which $\mathbf{F} \in \mathbb{C}^{M \times T}$ is a linear precoding matrix, such as Maximum-Ratio transmission (MRT) or ZF, and $\mathbf{x} \in \mathbb{C}^{T \times 1}$ is the signal vector. The signal received by the k-th terminal antenna can also be expanded in terms of the desired signal, interference, and noise, similar to Equation (3),

$$\hat{\mathrm{y}}_{\mathrm{dl}_k} = \sqrt{q}\mathbf{g}_{\mathrm{dl}_k} \cdot \mathbf{F} \cdot \mathbf{x} + n_k = \sqrt{q}\mathbf{g}_{\mathrm{dl}_k} \cdot \sum_{i=1}^{T} \mathbf{f}_i \mathrm{x}_i + n_k,$$

$$= \underbrace{\sqrt{q}\mathbf{g}_{\mathrm{dl}_k} \cdot \mathbf{f}_k \cdot \mathrm{x}_k}_{\text{desired signal}} + \underbrace{\sqrt{q} \sum_{j=1, j \neq k}^{M} \mathbf{g}_{\mathbf{dl}k} \cdot \mathbf{f}_j \cdot \mathrm{x}_j}_{\text{interference}} + \underbrace{n_k}_{\text{noise}}, \quad (5)$$

where $\mathbf{g}_{\mathrm{dl}_k}$ is the k-th row vector of the DL matrix $\mathbf{G_{dl}}$, $\mathbf{f}_i$, $\mathbf{f}_k$ and $\mathbf{f}_j$ are respectively the i-th, k-th and j-th column vectors of the precoding matrix $\mathbf{F}$, x_i and x_j are respectively i-th and j-th symbols sent by the BS, n_k is noise perceived by the k-th terminal antenna in the DL.

If we assume a Time-Division Duplex (TDD) mode, then $\mathbf{G_{dl}}$ can be calculated as the transpose of the uplink channel matrix $\mathbf{G_{ul}}$ (BJÖRNSON *et al.*, 2016). Typically, massive MIMO systems employ a TDD scheme in order to acquire CSI since it demands only UL pilot signals. Frequency-Division Duplex (FDD) schemes can also be employed by massive MIMO systems, however, when employing FDD schemes the BS must acquire CSI for UL and DL channels as they are not reciprocal to each other (LU *et al.*, 2014). In fact, the time required to transmit DL pilot symbols is proportional to the number of BS antennas (LU *et al.*, 2014). Thus, in FDD mode, the system is not scalable regarding the number of BS antennas. In this work, only the TDD mode operation is considered.

The instantaneous achievable rate C_{ul} for the UL can be calculated (in $[\mathrm{bits/s}]$) based on knowledge of the Shannon theory as:

$$C_{ul} = \log_2 \left[\det \left(\mathbf{I}_M + \frac{\rho_{ul}}{T} \mathbf{G_{ul}} \mathbf{G_{ul}}^{\mathrm{H}} \right) \right] [\mathrm{bits/s}], \quad (6)$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix and $(\cdot)^{\mathrm{H}}$ is the *conjugate-transpose* operation. For the DL case, the Equation (6) becomes:

$$C_{dl} = \log_2 \left[\det \left(\mathbf{I}_T + \frac{q}{M} \mathbf{G_{ul}}^{\mathrm{H}} \mathbf{G_{ul}} \right) \right], \quad (7)$$

where $\mathbf{I}_T$ is the $T \times T$ identity matrix. In both Equation (6) and Equation (7), the transmit power is constant and does not depend on the number of antennas, which is expressed in the normalization by M and T (MARZETTA *et al.*, 2016). To be valid, the Equations (6) and (7) assume that the additive receiver noise is complex Gaussian. Assuming the elements of the

channel matrices of DL and UL to be independent, identically distributed (i.i.d) zero-mean complex Gaussian, unit variance random variables and also that the Signal-to-Noise Ratio (SNR) is high enough, then random matrix theory yields that:

$$C \propto \min(M,T) \log_2(\rho_{snr}), \tag{8}$$

where ρ_{snr} is the SNR at the receiver side (MARZETTA, 2015).

The lower and upper bounds on the channel capacity are derived in (RUSEK *et al.*, 2013) and expressed as:

$$\log_2(1+\rho M) \leq C \leq \min(M,T) \log_2\left(1+\frac{\rho \max(M,T)}{T}\right). \tag{9}$$

where the scalar ρ is the SNR of the link, proportional to the transmitted power divided by the noise-variance (RUSEK *et al.*, 2013).

Theoretically, the channel capacity can be linearly scaled by increasing T and M. However, scaling point-to-point MIMO systems is not practical beyond a certain number of antennas on both sides. For instance, the 802.11ac standard sets a limit of 8 for both T and M. According to (MARZETTA, 2015), scaling point-to-point MIMO beyond 8×8 is not easy for a variety of reasons:

1. Line-of-sight conditions present a particular challenge because, for compact arrays, the channel matrix has a minimum rank of one, allowing only one data stream.

2. Increasing the number of antennas necessitates spending proportional amounts of time on training.

3. Near the edge of the cell, Signal-to-Interference-plus-Noise ratio (SINR) are typically low, and multiplexing gains fall short of the promised $\min(M,T)$.

4. The terminal equipment is complex, requiring separate electronic circuits for each antenna and advanced digital processing to separate the data streams.

5. To achieve performance close to the Shannon limit, both the BS and the user terminal must perform complex signal processing.

A slightly better alternative to the point-to-point MIMO is the multi-user MIMO system.

2.3.2 Multi-User MIMO

A simple representation of the MU-MIMO system is shown in Figure 2. The main idea of MU-MIMO is to split the T terminal antennas of the point-to-point MIMO system into T single-antenna devices scattered within the cell.

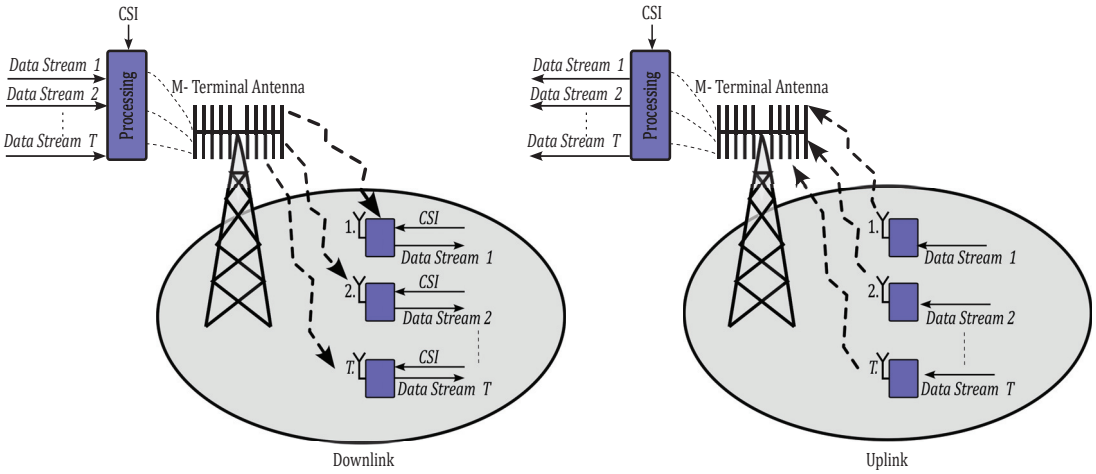


Figure 2 – Illustrative representation of the MU-MIMO system

This approach is better than the point-to-point MIMO for two main reasons: First, only the BS requires more expensive hardware, while the terminal sides can be equipped with relatively cheap antennas and circuits (LU *et al.*, 2014). Second, Line-of-sight conditions are no longer a significant issue (MARZETTA, 2015). This is because, in MU-MIMO systems, the terminals are usually separated by multiple wavelengths of distance, which renders their channel matrices also to be uncorrelated, allowing multiple data streams to be sent from multiple terminals (BJÖRNSON *et al.*, 2017b). Similar to the point-to-point MIMO, the UL channel capacity C^{ul} in the MU-MIMO is expressed (in $[\text{bits/s}]$) as :

$$C^{ul} = \log_2 \left[\det \left(\mathbf{I}_T + \rho_{ul}\mathbf{G_{ul}}\mathbf{G_{ul}}^{\mathrm{H}}\right)\right] [\text{bits/s}]. \tag{10}$$

However, the transmit power is T times higher than in the point-to-point MIMO model because there is no power sharing among the antennas (MARZETTA *et al.*, 2016). On the other hand, the DL capacity C^{dl} is computed as convex optimization problem as:

$$\begin{aligned} C^{dl} = \max_{\boldsymbol{\theta}} \quad & \log_2 \left[\det \left(\mathbf{I}_M + q\mathbf{G_{ul}}\mathbf{D}_{\boldsymbol{\theta}}\mathbf{G_{ul}}^{\mathrm{H}}\right)\right], \\ \text{subject to} \quad & \mathbf{b}^{\mathrm{T}}\boldsymbol{\theta} = 1, \\ & \boldsymbol{\theta} \geq 0, \end{aligned} \tag{11}$$

where $\boldsymbol{\theta}$ is a $M \times 1$ power allocation vector, $\mathbf{D}_{\boldsymbol{\theta}} = \text{diag}\left(\boldsymbol{\theta}\right)$ is a $M \times M$ positive diagonal matrix containing the power allocations elements of the vector $\boldsymbol{\theta}$, as its diagonal elements, $\mathbf{b}$ is a

$M \times 1$ vector of ones and $(\cdot)^{\mathrm{T}}$ is the *transpose* operation (LU *et al.*, 2014; MARZETTA *et al.*, 2016; MARZETTA, 2015).

One well-known way of allocating the diagonal elements of $\mathbf{D}_{\boldsymbol{\theta}}$ is through the *water-filling* algorithm. This power allocation algorithm consists in increasing the transmission power for streams that have better channel conditions, and it can be done either per user or jointly for all users in order to maximize overall capacity at the cost of an unfair power allocation (CRÂŞMARIU *et al.*, 2016).

However, to achieve the spectral efficiency showed in Equations (10) and (11) requires a more complex signal processing at both the BS and user terminal ends (the algorithm dirty paper coding/decoding) (MARZETTA, 2015). Also, the MU-MIMO system assumes that the downlink channel is known by both the BS and terminal devices, which is not practical as it requires that pilot signals are sent in both directions (MARZETTA *et al.*, 2016). Therefore, the multi-user MIMO as originally conceived is also not scalable.

2.3.3 Multiuser MIMO vs Massive MIMO

According to (BJÖRNSON, 2017), there are a few differences between MU-MIMO and massive MIMO that in the end turns the massive MIMO a much superior approach such as: **a**) In massive MIMO systems only the BS acquires CSI and simple linear signal processing is used on both the UL and the DL; **b**) In massive MIMO, the number of BS antennas is usually much higher, but not necessarily higher, than the number of users in the cell; **c**) In the MU-MIMO, channel estimation is done mainly based on codebooks with a set of predefined angular beams. In contrast, in massive MIMO, the channel is estimated based on UL pilots. An illustration of the massive MIMO system is shown in Figure 3.

The scalability of massive MIMO concerning the number of BS antennas allows the *channel hardening* effect to take place. The benefits of massive MIMO are also perceived from the single antenna terminal side, as the effective scalar channel behaves like AWGN. Therefore standard AWGN channel coding and modulation techniques are effective in most cases (MARZETTA *et al.*, 2016). For the BS to perform multiplexing and de-multiplexing of signal, massive MIMO depends on the acquisition of measured CSI from uplink pilots.

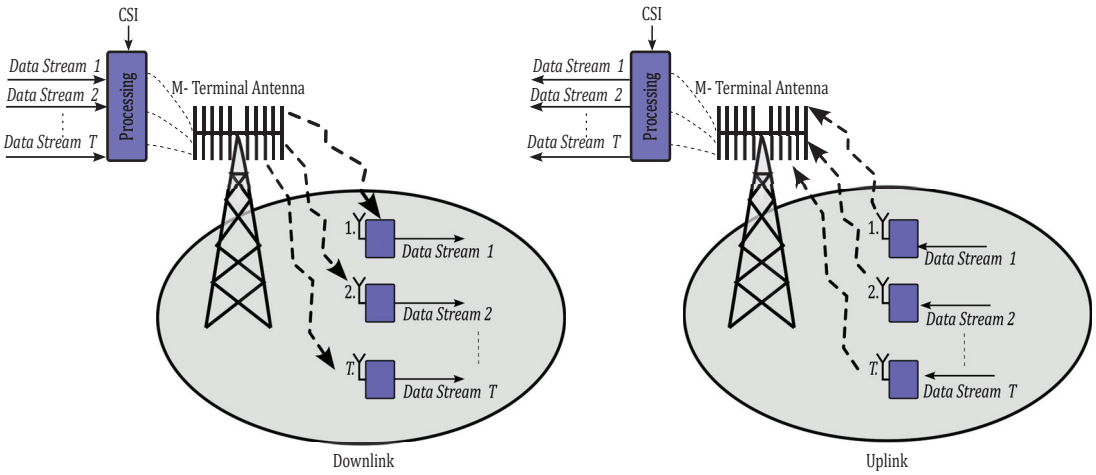


Figure 3 – Illustrative representation of the massive MIMO system

2.3.4 Coherence Interval

The coherence interval is a key concept in the context of B5G, mMTC, and massive MIMO systems. It can be defined as the time-frequency block over which a wireless channel remains constant or coherent and is a measure of the stability of the channel. This stability is crucial for ensuring reliable and efficient communications, as it allows for accurate demodulation and decoding of the signal. Also, the coherence interval of a wireless channel determines the maximum length of a block of data that can be transmitted without encountering significant interference or errors.

The length of the coherence interval depends on various factors, such as the mobility of the transmitter and receiver, the environment in which the communication is taking place, and the signal's frequency band. In general, a longer coherence interval indicates a more stable channel, allowing longer data blocks to be transmitted without errors. For example, signals in higher frequency bands, such as mmWave bands, tend to have shorter coherence intervals than those in lower frequency bands. Additionally, channel conditions, such as multi-path fading and shadowing, can also shorten the coherence interval.

The coherence interval can be mathematically represented as a combination of the coherence time, T_c, and the coherence bandwidth, B_c. T_c refers to the time duration for which a wireless channel can be considered time-invariant, while B_c denotes the range of frequencies where a constant value can approximate the magnitude of the channel's frequency response.

The coherence time, T_c, can be approximated as the amount of time that a moving device takes to travel half a wavelength of the considered signal:

$$T_c = \frac{\lambda}{2.\nu} \ [\mathrm{s}], \tag{12}$$

where λ is the wavelength and ν is the device's velocity (MARZETTA *et al.*, 2016).

The relationship between the wireless channel, represented by $g(t)$, the input signal $x(t)$, and the output signal $y(t)$ over a time interval T_c, is defined by the impulse response of the wireless channel:

$$y(t) = \int_{-\infty}^{\infty} g(\tau)x(t-\tau)d\tau, \tag{13}$$

and the channel's frequency response given by:

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-i2\pi ft}dt. \tag{14}$$

The magnitude of the channel's frequency response,$|G(f)|$, is, in general, frequency-dependent. However, it can be approximated by a constant value over a B_c frequency interval (MARZETTA *et al.*, 2016). The value of B_c can be approximated as:

$$B_c = \frac{c}{|d_1 - d_2|} \ [\text{Hz}], \tag{15}$$

where c is the speed of the light, and $|d_1 - d_2|$ is the maximum length difference between different propagation paths from the transmitter to the receiver (MARZETTA *et al.*, 2016). According to Marzetta *et al.* (2016, p. 22): "As a first-order approximation, $|d_1 - d_2|/c$ is equal to the delay spread of the channel, and $g(t)$ is time-limited to $|d_1 - d_2|/c$ seconds."

A coherence interval, depicted in Figure 4, is a time-frequency space characterized by a time span of T_c seconds and a bandwidth of B_c Hz, which is the maximum time-frequency space where the wireless channel's effects can be described by the multiplication of a complex-valued gain Φ. The magnitude of $|\Phi|$ represents the scaling of the waveform's envelope, and $arg(\Phi)$ represents a shift in its phase (MARZETTA *et al.*, 2016).

Finally, the sampling theorem states that any segment of a waveform spanning D seconds with energy limited to a B Hz wide frequency interval can be described in terms of BD complex-valued samples taken at each $1/B$ seconds. This fact yields that T_cB_c complex-valued samples are necessary to define a waveform fitting a coherence interval (MARZETTA *et al.*, 2016). The length of the coherence interval is then defined as:

$$\tau_c = T_cB_c \ [\text{samples}]. \tag{16}$$

Some estimates of T_c, B_c, and τ_p values are given in Table 2 for some different propagation scenarios at a carrier frequency of 2 GHz.

Considering a TDD scheme, the τ_c samples of the coherence interval naturally divide themselves into UL and DL subintervals since only one end of the link transmit a signal each time

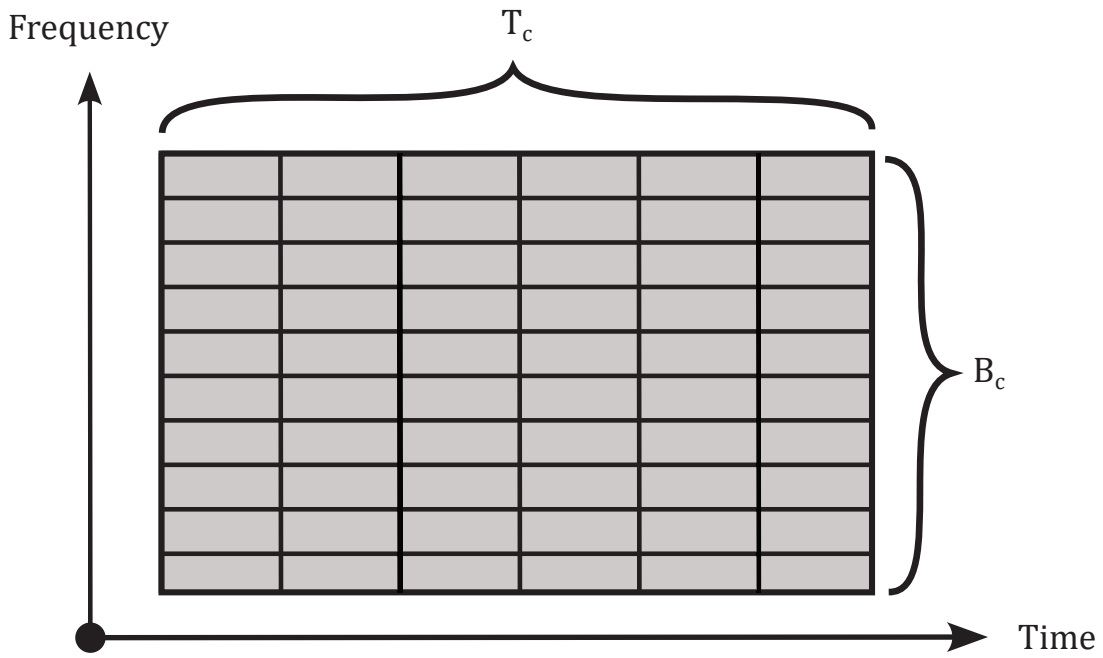


Figure 4 – Illustrative representation of the coherence interval

	Indoors $\lvert d_1 - d_2 \rvert = 30$ meters	Outdoors $\lvert d_1 - d_2 \rvert = 1000$ meters
Pedestrian $\nu = 1.5$ m/s (5.4 km/h)	$B_c = 10$ MHz $T_c = 50$ ms $\tau_c = 500{,}000$ samples	$B_c = 300$ kHz $T_c = 50$ ms $\tau_c = 15{,}000$ samples
Vehicular $\nu = 30$ m/s (108 km/h)	N/A	$B_c = 300$ kHz $T_c = 2.5$ ms $\tau_c = 750$ samples

Table 2 – First-order estimates of the coherence time T_c, coherence bandwidth B_c, and sample length of the coherence interval, τ_c, for some different propagation scenarios, at a carrier frequency of 2 GHz ($\lambda = 15$ cm).

Source: (MARZETTA *et al.*, 2016, p 23)

in the TDD mode. The UL and DL subintervals are then subdivided into UL and DL subintervals spent on uplink pilots and payload data. Hence, τ_c can be represented as a sum of subintervals as follows:

$$\tau_c = \tau_{ul} + \tau_{ul,p} + \tau_{dl} + \tau_{dl,p} \ \ [\text{samples}], \tag{17}$$

where τ_{ul} represents the number of samples per coherence interval spent on the transmission of UL payload data, $\tau_{ul,p}$ the number of samples per coherence interval spent on UL pilots, τ_{dl} the number of samples used for transmission of DL payload data, and $\tau_{dl,p}$ the number of samples allocated for DL pilots. Since the use of TDD mode allows the channel estimation of the DL by simply taking the conjugate-transpose of the UL channel, then in practice the $\tau_{dl,p}$ samples can be allocated for the DL payload data τ_{dl} and τ_c will constituted of only 3 subintervals:

$$\tau_c = \tau_{ul} + \tau_p + \tau_{dl} \ \ [\text{samples}], \tag{18}$$

in which the $(\cdot)_{ul}$ subscript is dropped from the subinterval $\tau_{up,p}$ to simplify the notation.

The shortage of coherence blocks is closely tied to the scarcity of pilot signals 2.3.5. Measuring the number of coherence blocks in each use case scenario is crucial for any B5G project design, especially in mMTC systems where the number of connected devices to the BS may exceed the number of available pilots. This work proposes three RA protocols to handle use case scenarios with limited pilot availability.

2.3.5 Pilot Signals

In the massive MIMO technology *channel estimation* is a key step. The BS estimates the channel gain of the terminals, receiving and decoding pilot sequence signals sent by them. For the BS to estimate the channels of T terminals, at least T mutually orthogonal pilot waveforms are needed. The orthogonality among the pilots is required to ensure that there is no interference between them. Hence, the BS keeps available a set of τ_p orthogonal pilots sequence of length τ_p to be shared among devices on the network, where $\tau_c \geq \tau_p \geq T$ (MARZETTA *et al.*, 2016). The pilot sequence length must be equal to or greater than the number of pilots for them to be orthogonal. Naturally, τ_p is the smallest possible pilot length capable of ensuring orthogonality within a set of τ_p pilots.

The BS uses the following procedure to estimate the channel gains of terminals using pilots:

When the k-th terminal wants to transmit a signal to the BS, it chooses one of the available pilot sequences, herein represented by the vector $\boldsymbol{\psi_k}$ with dimension $\tau_p \times 1$. Collectively, all pilot sequences chosen by T terminals is then represented by a $\tau_p \times T$ unitary matrix $\boldsymbol{\Psi} = [\boldsymbol{\psi_1}, \boldsymbol{\psi_2}, ..., \boldsymbol{\psi_k}]$ with $\tau_p \geq T$ and

$$\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi} = \mathbf{I}_T. \tag{19}$$

Together, all T terminals transmit a $T \times \tau_p$ signal,

$$\mathbf{X}_p = \sqrt{\tau_p}\boldsymbol{\Psi}^{\mathrm{H}}, \tag{20}$$

which, according to Marzetta *et al.* (2016, p. 46), "is normalized so that each terminal expends a total energy that is equal to the duration of the pilot sequence",

$$\tau_p \boldsymbol{\psi}_k^{\mathrm{H}} \boldsymbol{\psi}_k = \tau_p. \tag{21}$$

The pilot signals sent by the UL channel is then received at the BS as a $M \times \tau_p$ matrix signal,

$$\mathbf{Y}_p = \sqrt{\rho_{ul}}\mathbf{G}_{ul}\mathbf{X}_p + \mathbf{N}_p = \sqrt{\tau_p \rho_{ul}}\mathbf{G}_{ul}\mathbf{\Psi}^{\mathrm{H}} + \mathbf{N}_p, \tag{22}$$

where $\mathbf{N}_p$ is the receiver $M \times \tau_p$ noise matrix, with entries i.i.d, following a complex normal distribution, $\mathcal{CN}(0, 1)$.

The BS then correlates $\mathbf{Y}_p$ with each of the T pilot sequences, yielding a noisy version of the channel matrix $\mathbf{G}_{ul}$:

$$\mathbf{Y}'_p = \mathbf{Y}_p\mathbf{\Psi} = \sqrt{\tau_p \rho_{ul}}\mathbf{G}_{ul}\mathbf{\Psi}^{\mathrm{H}}\mathbf{\Psi} + \mathbf{N}_p\mathbf{\Psi} = \sqrt{\tau_p \rho_{ul}}\mathbf{G}_{ul} + \mathbf{N}'_p, \tag{23}$$

where $\mathbf{N}'_p = \mathbf{N}_p\mathbf{\Psi}$ is $M \times T$ noisy matrix, whose entries are also i.i.d with distribution $\mathcal{CN}(0, 1)$ (MARZETTA *et al.*, 2016). The entries of $\mathbf{G}_{ul}$ are then estimated by employing a channel estimator. For instance, under the assumption of independent Rayleigh fading, the elements of the channel matrix and the noise matrix are statistically independent, and the Minimum Mean Square Error (MMSE) method can be employed to estimate the (m,k)-th component of $\mathbf{G}_{ul}$ as follows:

$$[\mathbf{Y}'_p]_{mk} = \sqrt{\tau_p \rho_{ul}}g_k^m + [\mathbf{N}'_p]_{mk}. \tag{24}$$

Assuming that the large-scale fading coefficients are known, so the prior distribution of g_k^m , $\mathcal{CN}(0, \beta_k)$, is also known. The MMSE estimator yields:

$$\hat{g}_k^m = \mathbb{E}[g_k^m|\mathbf{Y}_p] = \mathbb{E}[g_k^m|\mathbf{Y}'_p] = \frac{\sqrt{\tau_p \rho_{ul}}\beta_k}{1 + \tau_p \rho_{ul}\beta_k}[\mathbf{N}'_p]_{mk}. \tag{25}$$

2.3.6 Pilot Contamination and Pilot Collision

One of the main benefits of massive MIMO is to allow simple and linear estimators such as MMSE and Least Squares (LS) to be employed by BS for acquiring CSI (BJÖRNSON *et al.*, 2017b). Researchers have shown that signal detection, and precoding can also be carried out using simple and linear signal processing approaches such as Matched-Filter (MF), ZF and Maximum Ratio (MR) when the number of BS antennas is large without a significant performance loss when compared to more complex non-linear methods (RUSEK *et al.*, 2013).

Ideally, the number of pilots should be equal to the number of devices. Nonetheless, since both time and frequency spectrum resources are limited, the number of possible orthogonal pilots is also limited. Hence, pilot reuse is frequently necessary either in the same or neighboring cells.

Currently, pilot contamination and pilot collision are the main impairments to massive MIMO scalability. Pilot contamination occurs due to the reuse of pilots in neighboring cells; it acts as an interference and does not disappear by increasing M. Pilot collisions occur due to the reuse of pilots within the same cell.

Nonetheless, as the number of mobile phones and mMTC devices is continuously increasing, the number of available pilots will eventually not be enough for the BS to serve all devices within the cell. This shortage of pilots gives rise to performance issues such as *pilot collisions* when two or more users choose the same pilot to try to access the BS resources. Since the BS cannot estimate the channel when a pilot collision occurs, none of the contenders can access the BS resources. This issue leads to a system bottleneck and makes the network unsuitable for high data transmission devices and URLLC. Therefore, establishing a RA policy is mandatory.

2.4 RANDOM ACCESS PROTOCOLS

Several RA protocols have already been proposed. A RA policy can be established mainly either as GF or GB handshake procedure. In a GB process, the device first has to be granted permission to access the system resources controlled by the BS before sending any payload data through an exclusive channel. On the other hand, in a GF process, the device is allowed to send its payload data without being granted an exclusive channel of the system. This section reviews the most common RA procedures proposed in the literature.

2.4.1 Grant-based Protocols

A common GB procedure is the LTE Random Access Channel (RACH) protocol, Figure 5, which has currently been applied to the narrowband IoT (NB-IoT) (BJÖRNSON *et al.*, 2017b). The LTE RACH can be either a contention-based random access or a non-contention-based random access. The non-contention-based random access procedure has three steps: First, the BS allocates random access preambles to all devices. Next, each device sends the assigned preamble to the BS. Finally, after receiving the preambles, the BS sends the Random Access Response (RAR) to all devices. Since each device has its own preamble, a collision in this procedure can be avoided.

The contention-based procedure, shown in Figure 6, has four steps: In the first step,

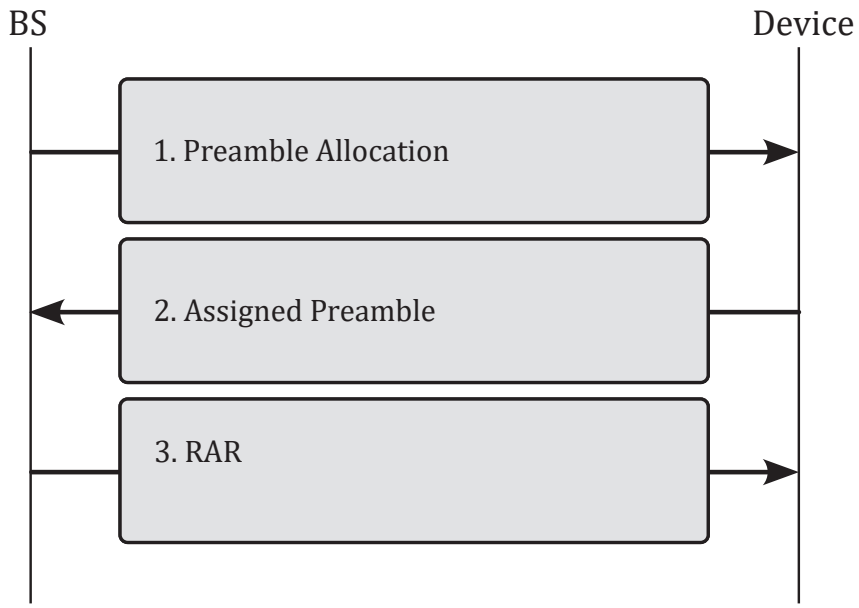


Figure 5 – Non-contention-based RACH Protocol for LTE systems

the devices send a preamble randomly chosen from a set of preambles made available by the BS. Then, the BS send a RAR to the devices whose preambles did not collide in the previous step. Next, After receiving a response to its preamble transmission, each device sends a Radio Resource Control (RRC) connection request to request resources for ensuing data transmission. If more than one device activated the preamble, all of these devices use the same resource to send their RRC connection request in Step 3, and the BS detects this collision. In the last step, the BS achieves the contention resolution in the DL, a complex procedure that may require all colliding devices to make a new access attempt after a random waiting period (WANG; MA, 2019).

Random Access GB protocols targeted to massive MIMO systems have also been proposed. In (BJÖRNSON *et al.*, 2017a), Figure 7, the SUCRe protocol is introduced. This protocol is a GB, 4-steps RA procedure whose idea is to grant access, to the BS resources, only to the strongest pilot contender each time. Firstly, the devices that wish to become active transmit during the UL a RA pilot sequence randomly chosen from the set of pilots provided by the BS. In the next step, the BS responds with a precoded pilot in the DL, with the precoding evaluated based on CSI acquired from each RA pilot received. Receiving this signal allows devices to assess the total signal strength of all contenders competing for the same RA pilot. Based on this information, each competitor decides whether or not to retransmit its pilot signal in the UL of step 3. The pilot is retransmitted only when the device's signal strength exceeds 50% of the entire sum of the signal strengths of all devices competing for the same specific RA pilot. Finally, in the last step, the BS assigns dedicated communication resources to the devices whose pilot signals did not collide in step 3. In numerical Monte-Carlo simulations, the SUCRe protocol has been

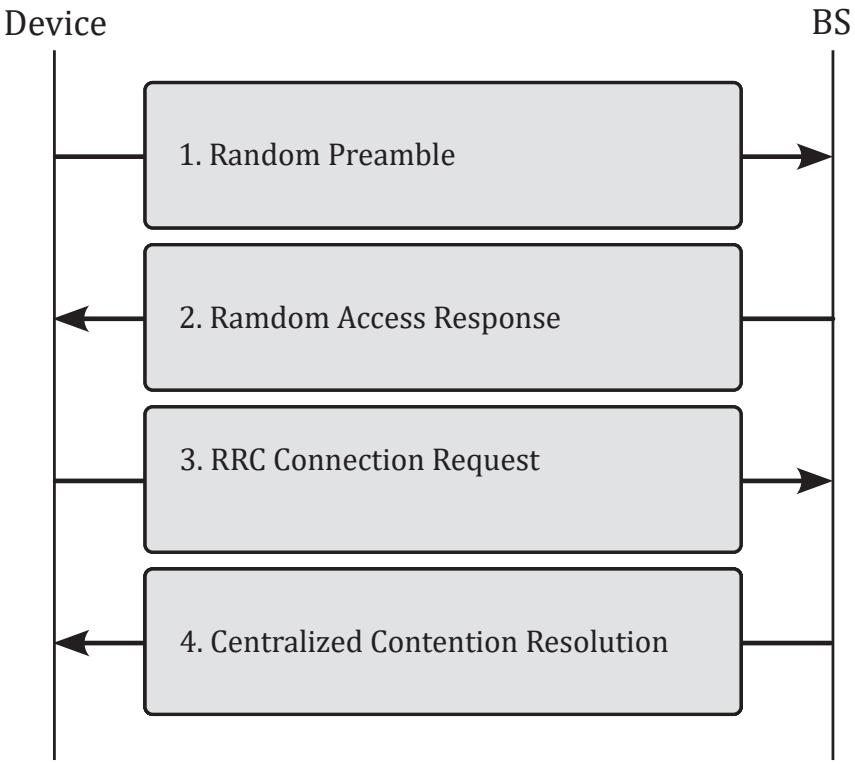


Figure 6 – Contention-based RACH Protocol for LTE systems

shown to be very successful, solving around 90% of all collisions. However, the channel gain correlates with the device's distance from the BS. Thus, in the RA stage, devices closer to the BS are favored over those farther away. Moreover, the protocol does not resolve false-negative pilot collisions in which the strongest device has a signal strength value below 50% of the total sum of the contesting devices' signal strengths. Furthermore, the 4-step procedure required for the BS to grant exclusive communications resources to the devices could be a performance bottleneck and a source of excessive delay and signaling overhead. Therefore, it is not the ideal choice for mMTC systems, where accessing devices usually have small data packets to transmit sporadically. Other works, such as (HAN *et al.*, 2017a; MARINELLO; ABRÃO, 2019; HAN *et al.*, 2017b; MARINELLO *et al.*, 2020), present proposals for optimization of the SUCRe protocol showing promising results. However, they are also GB protocols that introduce extra complexity or overhead compared to the SUCRe protocol. To support new use cases, B5G RA schemes should achieve high scalability under latency and reliability constraints. For this purpose, grant-free (GF) RA protocols have gained increasing interest, as they can drastically reduce control signaling for connection establishment (CHEN *et al.*, 2021).

A grant-based protocol is better suited to eMBB applications that require high data rates and reliable connections. However, when dealing with massive machine-type devices, a four-step handshake procedure may result in excess of control signaling overhead. Given that most use

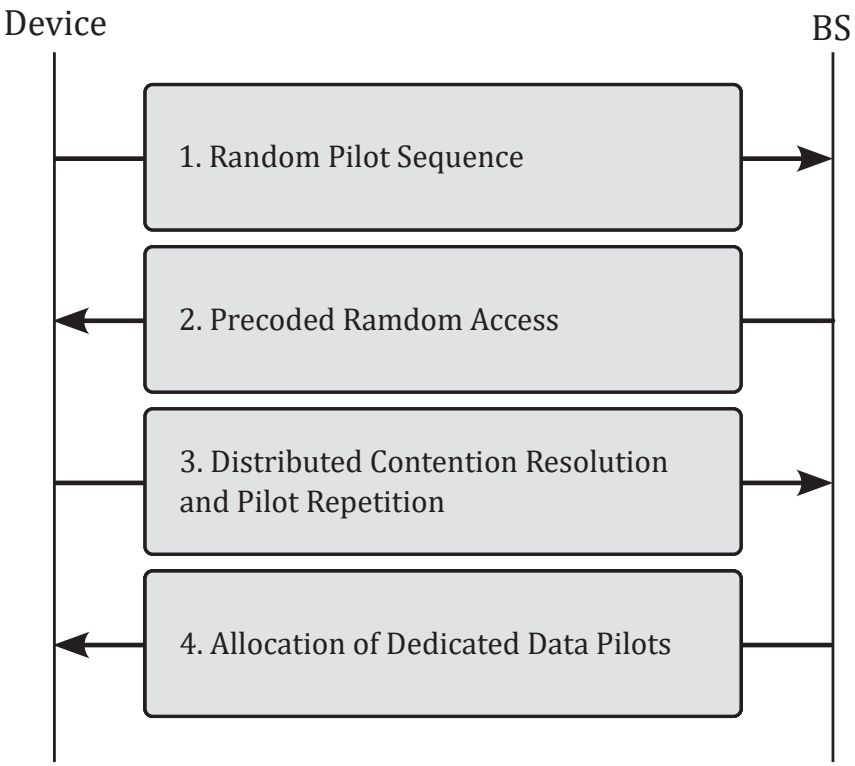


Figure 7 – Illustrative representation of the SUCRe Protocol

cases of mMTC systems are in the form of small sensors and actuators that transmit small data packets, a GF RA protocol is a better approach.

2.4.2 Grant-free Protocols

Classical GF RA approaches are the contention resolution diversity slotted ALOHA (CASINI *et al.*, 2007) and irregular repetition slotted ALOHA (LIVA, 2011) protocols. The idea behind those schemes is to repeat the transmission of data packets in several randomly chosen slots. When a device uniquely selects a slot, its payload is successfully decoded, and its interference in other slots is canceled via Successive interference cancellation (SIC), increasing the occurrence of other non-colliding slots. Although these protocols achieve excellent results, they have the disadvantages of requiring packet retransmissions, overhead for side information signaling, increased complexity for SIC evaluation, and the possibility of propagation errors.

Grant-free random access protocols can also take advantage of the massive MIMO properties. In (BAI *et al.*, 2021), a GF random-access protocol that minimizes signaling overhead and access delay, and exploits the features of massive MIMO to handle preamble collisions, is proposed. The main novelty introduced by (BAI *et al.*, 2021) is an algorithm capable of decoding up to three collided signals via the following steps: First, the BS broadcasts system information

to the devices, allowing them to estimate their channel gains and use Fractional Power Control (FPC) to compensate for path loss. Following that, devices wishing to connect to the network randomly select a preamble from a predetermined list assigned to the corresponding cell and transmit the preamble, followed by data, to the BS. Next, using the energy detector, the BS detects preambles in use and estimates the channel vector corresponding to each detected preamble. If multiple devices select the same preamble, the corresponding channel estimate transforms into an estimate of the superposition of the multiple devices' channels. The BS then employs an Automatic Modulation Classification (AMC) procedure based on elementary High-Order Cumulants (HOCs) to detect the presence of collisions through a given threshold. If there is no collision, the decoding phase ends. If the BS detects a collision, the data sequence is forwarded to the next stage, where the BS employs a Second-Order Statistic of the Eigenvalues (SORTE) approach to determine the number of devices that experience preamble collisions. Finally, each signal layer must be decoded using an appropriate SIC strategy. For each superimposed signal, the BS decodes the two strongest signal layers and attempts to decode the third signal layer only when the estimated number of colliding devices is greater than twice the estimated number of preambles picked by multiple devices. In (DING *et al.*, 2019), analytic expressions of success probability of the GF RA applied in massive MIMO networks for conjugate and zero-forcing beamforming techniques are derived, showing that GF protocols are an attractive RA technique with low signaling overhead in massive MIMO networks. It is also shown that massive MIMO networks employing GF RA protocols could simultaneously accommodate a number of RA users, which is multiple times the number of RA channels, with almost 100% success probability. GF RA protocols are also applied to non-orthogonal multiple access (NOMA) and IoT technologies (SHAHAB *et al.*, 2020). Other works, such as (BELLO *et al.*, 2014) and (SHARMA; WANG, 2019) that apply ML algorithms for pilot collision resolution, will be discussed in the next section. Figure 8 illustrates a generic diagram of a 2-step GF procedure, in which a device sends the pilots signal followed by the data payload to the BS and receives an acknowledgment signal (ACK) if the BS successfully decodes the device's signal.

2.5 MACHINE LEARNING

The term *Machine learning* (ML) refers to a part of AI and computer science that studies algorithms capable of imitating the way humans learn (IBM, 2023). In the context of wireless communications, including B5G , ML based techniques have been finding a plethora of new

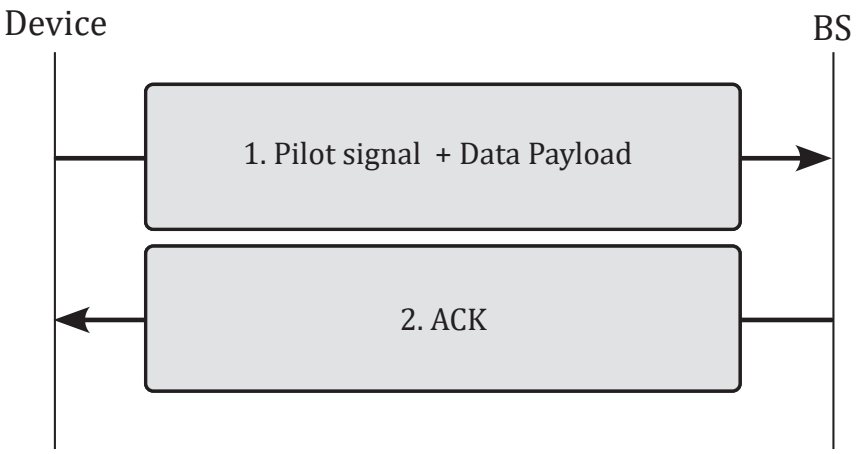


Figure 8 – 2-step *grant-free* protocol

possible applications such as in (KOLODZIEJ *et al.*, 2020), (MUSTAFA *et al.*, 2021), and others. There are three main types of ML algorithms: **a**) Supervised Learning; **b**) Unsupervised Learning; and **c**) Reinforcement Learning. Both supervised learning and unsupervised learning need training data. However, while supervised learning requires labeled training data, unsupervised learning is able to navigate through data and find groups of data that might form a specific data category. Reinforcement learning, however, does not require any training data and learns from the environment by expressing a behavior and being rewarded for it.

2.5.1 Supervised Learning

A supervised learning (SLe) model is a computational algorithm that predicts outcomes based on labeled training data. To do this, the SLe model requires a set of labeled training data where each data point consists of an input and its corresponding label or desired output. The goal of the model is to learn a function that maps inputs to their corresponding outputs.

Before training an SLe model, preparation of the training data is necessary. This may involve splitting the data into training and test sets in a predetermined proportion, such as 80% for training and 20% for testing. Preprocessing the data, such as normalizing it to values between 0 and 1, may also be required.

A classic example of labeled training data is the *iris* dataset. This dataset is a table where each row represents an iris flower, and the columns present information taken from each individual flower. The information includes the species (labels) and features such as the dimensions of their botanical parts, sepal, and petal, in centimeters. The *iris* dataset has four features: "Sepal length", "Sepal width", "Petal length", and "Petal width", and three possible species labels: "Versicolor", "Setosa", and "Virginica".

In the example of the *iris* dataset, the features of each data item can be used as a 4-dimensional input vector of a training model so that each input vector corresponds only to a single output class. The model's output can be numerically modeled through the encoding method *one-hot* for classification tasks, where each class is arbitrarily modeled as a vector with a dimension equal to the number of classes. In the *one-hot* model, each element of the output class is either 0 or 1, only a single element of the vector output is equal to 1, and all others are equal to 0. Hence, each input vector $\mathbf{input} = [\text{"Sepal length"}, \text{"Sepal width"}, \text{"Petal length"}, \text{"Petal width"}]^{\mathrm{T}}$ corresponds to one of three possible output vector classes:

$$\mathbf{output} = \left(\underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\text{Versicolor}}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\text{Setosa}}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\text{Virginica}} \right). \tag{26}$$

This example can be generalized for any classification task with known desired output classes, regardless of the number of features or classes. In general, SLe methods can be used to estimate any function of type $f : X \to Y$, where $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$, in which m and n are respectively the input and output vectors' dimensions.

The most common types of SLe based methods are *Linear Regression*, a technique that, by fitting a linear equation to the observed data, seeks to predict the value of one or more continuous target variables d given the value of a D-dimensional vector $\mathbf{x}_d$ of input variables (BISHOP, 2006). *Logistic Regression*, a method used for binary classification tasks in which a logistic curve is fit to the observed data (SHALEV-SHWARTZ; BEN-DAVID, 2014). *Decision Tree*, a tree-like hierarchical model that, by traveling from a root node of a tree to a leaf (where the label is found), can be used to predict the label associated with a given input (SHALEV-SHWARTZ; BEN-DAVID, 2014). *Random Forest*, a collection of decision trees whose predictions are combined to improve the overall performance and reduce the variance (SHALEV-SHWARTZ; BEN-DAVID, 2014). *Support Vector Machine (SVM)*, a method that works by finding a hyperplane that maximally separates the different classes of data (SHALEV-SHWARTZ; BEN-DAVID, 2014). Finally, *Neural Networks*, a network of artificial neurons that are organized in layers and trained for learning the correct output for a given input data (SILVA *et al.*, 2016).

There are many examples of SLe applications in the context of wireless communications. For instance, the SLe version of *deep learning* has been employed to solve various issues in

B5G systems such as in (CHUN *et al.*, 2019) where a deep learning based approach for channel estimation in massive MIMO systems is proposed. The suggested method is applied to cases when the pilot length is smaller than the number of transmit antennas achieving better results than traditional approaches. Another example is as described in (WEN *et al.*, 2018) where deep learning is used as a tool for minimizing overhead when acquiring downlink CSI for massive MIMO systems employing FDD mode, which is not practical using conventional methods. Other ML approaches, including *Decision-Tree*,*Random Forests*, *Space-Vector-Machine*,*Convolutional Neural Networks* and *Recurrent Neural Network* are also applied in various issues within B5G systems (REKKAS *et al.*, 2021).

2.5.2 Unsupervised Learning

Unlike a SLe model, an unsupervised learning (ULe) model does not rely on labeled data items. Instead, ULe algorithms aim to discover patterns in the data collection. An important ULe application is clustering, which is finding groups of data within the dataset. The clustering model, however, is not able to provide specific labels for those groups. For instance, a clustering model presented with an unlabeled version of the *iris* dataset is capable of finding 3 different groups. However, only a specialist can define the specific class of each group as "Versicolor", "Setosa" or "Virginica". One of the most popular clustering approaches is the *K-Means* method, a centroid-based algorithm where each data group is associated with a centroid. The algorithm iteratively updates the centroid and the membership of data points to the clusters until convergence (SHALEV-SHWARTZ; BEN-DAVID, 2014).

Other types of ULe applications are dimensionality reduction and anomaly detection. The goal of dimensionality reduction is to reduce the number of variables in data while retaining as much information as possible. This can be useful for visualizing high-dimensional data or improving the performance of supervised learning algorithms. Some popular dimensionality reduction methods include Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Anomaly detection, also known as outlier detection, aims to identify data items that do not conform to the normal pattern in the data.

Unsupervised learning-based algorithms, such as *K-Means*, Self-Organizing Maps (SOM), and unsupervised Long Short-Term Memory (LSTM), have been effectively applied in various scenarios, including power allocation, fault detection, fault management, and channel estimation (REKKAS *et al.*, 2021). In (CUI *et al.*, 2018), the *K-Means* algorithm is utilized to

solve the sum-rate maximization problem in a mmWave-NOMA system where the distribution of users' locations is modeled using a Poisson cluster, resulting in improved performance compared to applying *K-Means* to Orthogonal multiple access (OMA) systems combined with mmWave technology. The sum-rate maximization problem for DL in mmWave-NOMA systems is also studied in (REN *et al.*, 2019). The authors propose an Expectation Maximization (EM) algorithm for user clustering in a fixed user model, demonstrating superiority over the same algorithm applied to mmWave-OMA systems. Other works such as (GóMEZ-ANDRADES *et al.*, 2016) and (FARSAD; GOLDSMITH, 2017) also propose unsupervised learning-based solutions, achieving good results.

2.5.3 Reinforcement Learning

Reinforcement learning (RLe) based methods differ from SLe and ULe methods, as they do not require data items for training. The model (an agent) learns by interacting with the environment and receiving rewards and penalties for its actions. The goal is to choose actions that will maximize the cumulative reward over time.

A classic RLe method is the Q-Learning, an algorithm based on the Q-function, which estimates the maximum expected cumulative reward for an agent following a particular policy (SUTTON; BARTO, 2018). The Q-Learning method operates by constructing a table (Q-table) that stores the cumulative reward $Q(s, a)$ for taking a certain action a if the agent is currently in state s and following the optimal policy from there on. The cumulative reward $Q(s, a)$ is updated as:

$$Q_{t+1}(s, a) = Q_t(s, a) + \kappa(r + \gamma * \max_a(Q(s_{t+1}, a_{t+1})) - Q_t(s, a)), \tag{27}$$

where κ is the learning rate, a value between 0 and 1 that determines the update rate, r is the immediate reward for taking action a in state s, γ is the discount factor, a value between 0 and 1, that controls the trade-off between short-term and long-term rewards. For example, if γ is close to 0, the agent will prioritize short-term rewards, while if γ is close to 1, the agent will prioritize long-term rewards. s_{t+1} is the next state after taking action a in state s. a_{t+1} is the next action taken in the next state s_{t+1}. $\max_a(Q(s_{t+1}, a_{t+1}))$ is the maximum action-value for the next state s_{t+1}. This is the maximum expected cumulative reward for all actions in the next state. Hence, the Q-table works as a map for the agent to find the action with the highest reward

in its current state. Other RLe based methods are *State–Action–Reward–State–Action (SARSA)*, *Temporal-Difference (TD)* learning, *Monte Carlo Methods*, and *deep reinforcement* learning (SUTTON; BARTO, 2018).

Due to the randomly ever-changing nature of wireless channels, RLe based methods are interesting approaches to apply in B5G and other wireless communications technologies as they can adapt to environmental changes in almost real-time. In (BELLO *et al.*, 2014), a GF RA RLe-based protocol is proposed as a pilot collision control method. The RLe algorithm used in (BELLO *et al.*, 2014) is the *Q-Learning*, in which each accessing device is an independent agent and the rewards are $+1$ if it selects a unique RA slot or -1 otherwise. As an improvement, (SHARMA; WANG, 2019) proposes a collaborative Q-learning RA scheme in which the negative rewards in the event of a collision are proportional to the congestion level of the chosen RA slot. However, the protocol assumes that the devices know the exact number of devices colliding by their chosen RA slot. The performance results are better than when compared with the independent Q-learning approach of (BELLO *et al.*, 2014). Nonetheless, both (BELLO *et al.*, 2014) and (SHARMA; WANG, 2019) do not take into account realistic effects like multipath fading, path loss, thermal noise, and ICI, besides assuming the devices know the exact congestion levels.

2.6 STATISTICAL INFERENCE

Classical Statistical Inference (SI) models such as the frequentist inference, Pearson Correlation, and the Bayesian method differ from ML models as they derive population inferences from a population sample while machine learning discovers generalizable predictive patterns (BZDOK *et al.*, 2018). In principle, many statistical and ML methods can be used for prediction and inference. However, statistical methods have long focused on inference, which is accomplished through developing and fitting a project-specific probability model (BZDOK *et al.*, 2018). The model enables us to calculate a quantitative measure of confidence that a discovered relationship describes a *true* effect that is unlikely to be caused by noise. Moreover, if sufficient data are available, we can explicitly validate assumptions (e.g., equal variance) and, if necessary, refine the specified model (BZDOK *et al.*, 2018).

Statistical learning is a process of data analysis in which a set of general characteristics of a population are inferred from a subset of the population. The process of statistical learning through Bayes' rule, (28), is called *Bayesian method* or *Bayesian inference* (HOFF, 2009).

The Bayesian method can be applied for binary classification tasks, as described in the following example. Suppose we want to classify fruits that pass at a given point of a conveyor belt in an industry. Also, suppose that these fruits are of only two types: "pears" and "apples". A class or *state of nature* C_j, with $j \in \{1,2\}$, is associated with each type of fruit. We let the state C_1 represent "pears," and C_2 represent "apples". We assume these fruits do not arrive at the observed point in any predictable order, so C_j is a random variable. Finally, we assume that all costs associated with incorrect classifications are equal. After some observations, it is possible to establish the *a priori* (prior) probability that the fruit is a pear $P(C_1)$ or an apple $P(C_2)$. The prior probability represents what is originally believed by the classifier before new evidence is introduced. An important clue to help in classifying the fruit passing at the observed point is to know which one has the highest associated prior probability $P(C_j)$. However, the probability $P(C_j)$ alone is not enough. If $P(C_1) > P(C_2)$, the classifier would always decide for C_1 if the only clue it had was $P(C_j)$. Complementing our classifier, we can take some measurements from the fruits using a sensor. For instance, we can measure their weights[1]. Letting x, a continuous random variable, be the weight of an observed fruit, the relationship between x and C_j can be established as the class-conditional *probility density function (PDF)* $p_x(x|C_j)$, which is the PDF of x given a class C_j. A probability density function is a function that describes the likelihood of a random variable taking on a particular value. One of the most important properties of a PDF is that the area under the curve must always equal 1, which represents the total probability of the random variable taking on any value within its domain. The definite integral of the PDF over a range of values gives the probability that the random variable falls within that range. A class-condition probability is the PDF of a conditional random variable, in which the condition is that the random variable belongs to a class C_j. The PDF term $p_x(x|C_j)$ is called the *likelihood* of C_j with respect to x. The likelihood term refers to the conditional probability or PDF of a particular measured value x, assuming that the x came from a specific scenario or class.

Figure 9 shows an illustration of hypothetical $p_x(x|C_1)$ and $p_x(x|C_2)$ PDFs overlapping each other. For the sake of this example, it is assumed that the PDF $p_x(x|C_1)$ is a gaussian curve with mean $\mu_1 = 2$ and standard deviation $\sigma_1 = 0.8$ and $p_x(x|C_2)$ is a gaussian curve with mean $\mu_2 = 3$ and standard deviation $\sigma_2 = 1$. The PDFs are also normalized, so the area under each one is equal to 1.

The prior probabilities $P(C_j)$ can be combined with the class-conditional PDF $p_x(x|C_j)$

[1] As this is a generic example, there is no specific weight unit mentioned in the body of the text. The weight unit can be of any type as long as it can be measured as a continuous real value.

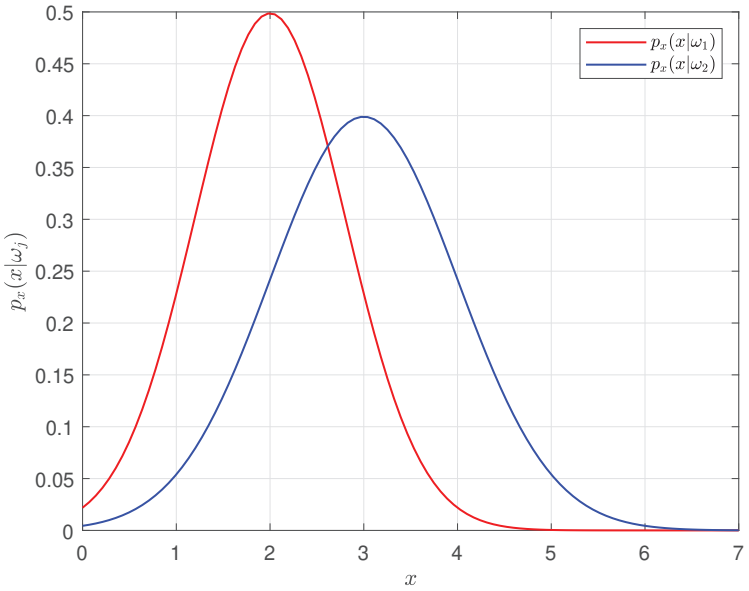


Figure 9 – Representative illustration of two overlapping gaussian curves. With $\mu_1 = 2$, $\mu_2 = 3$, $\sigma_1 = 0.8$ and $\sigma_2 = 1$.

through the Bayes' rule to calculate the *a posteriori* (posterior) probability $P(C_j|x)$ in Equation (28). Where $P(C_j|x)$ is the probability of the fruit being in a state of nature C_j given that x is its measured weight (DUDA *et al.*, 2000). A posterior probability calculated through the Bayes' rule expresses the likelihood of an event given the evidence that has been observed. In this example, the event is belonging to a class C_j given as evidence, the random variable x.

$$P(C_j|x) = \frac{p_x(x|C_j)P(C_j)}{p_x(x)}. \tag{28}$$

The Equation (29) presents the PDF $p_x(x)$ called *evidence*. The evidence $p_x(x)$ is the PDF of x being measured among all considered J classes. In this example, $J = 2$. (DUDA *et al.*, 2000).

$$p_x(x) = \sum_{j=1}^{J} p_x(x|C_j)P(C_j). \tag{29}$$

What Bayes' rule says is that given a feature value x, the prior probability $P(C_j)$ can be converted into a posterior probability $P(C_j|x)$ of the *state of nature* being C_j given a measured feature x. The evidence factor $p_x(x)$ is only a scale factor that guarantees that posterior probabilities sum to 1 (DUDA *et al.*, 2000).

Figure 10 depicts the $P(C_1|x)$ and $P(C_2|x)$ posterior probabilities as a functions of x and assumes prior probabilities of $P(C_1) = 0.3$ and $P(C_2) = 0.7$.

The decision procedure indicates which of the states of nature, C_1 or C_2 is more likely to be the true one. The natural choice is to take the one with the highest posterior probability

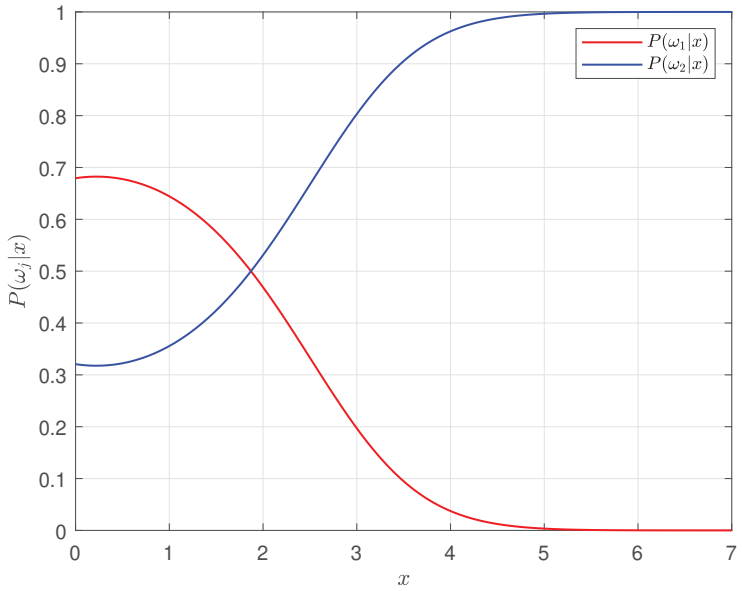


Figure 10 – Representative illustration of $P(C_1|x)$ and of $P(C_2|x)$ as functions of x. With $P(C_1) = 0.3$ and $P(C_2) = 0.7$.

for any particular observation of x. For instance, if $P(C_1|x) > P(C_2|x)$, the classifier decides that the true *state of nature* for a measured feature x is C_1 and C_2 if $P(C_2|x) > P(C_1|x)$. What justifies this choice is that it minimizes the joint probability $P(error,x)$ of a error occurring in the decision procedure given a value x:

$$P(error,x) = \begin{cases} P(C_1|x) & \text{if decide for } C_2 \\ P(C_2|x) & \text{if decide for } C_1 \end{cases} \tag{30}$$

Since there are only two classes and $P(C_1|x)+P(C_2|x) = 1$, the probability of error for deciding for C_1 will always be $P(C_2|x)$ and vice versa. This fact justifies the decision procedure. In this example, the Bayesian classifier would decide for C_1 (the fruit is classified as a pear) if $x < 1.909$ and for C_2 (the fruit is classified as an apple) otherwise.

A generalization of this example involves incorporating multiple features, a number of classes larger than just two, and different types of actions beyond classification tasks. This also involves the introduction of a loss function to quantify the cost of an action. The multiple features can be represented by a random vector $\mathbf{x}_d^R \in \mathbb{R}^d$, where d is the number of features. The set of J classes is represented as $\{C_1,..,C_J\}$ and the set of C possible actions as $\{a_1,..,a_C\}$. The loss function $\lambda(a_i|C_j)$ represents the loss incurred when taking action a_i when the true class is C_j. With these elements, Bayes' Rule (Equation (28)) can be written in a more general form:

$$P(C_j|\mathbf{x}_d^R) = \frac{p_x(\mathbf{x}_d^R|C_j)P(C_j)}{p_x(\mathbf{x}_d^R)}. \tag{31}$$

where $p_x(\mathbf{x}_d^R|C_j)$ is the class-conditional PDF of $\mathbf{x}_d^R$, with the probability density function for $\mathbf{x}_d^R$ given C_j, the true class. As before, $P(C_j)$ describes the prior probability that the state of nature C_j is true, $P(C_j|\mathbf{x}_d^R)$ is the posterior probability. Equation (29) is rewritten as:

$$p_x(\mathbf{x}_d^R) = \sum_{j=1}^{J} p_x(\mathbf{x}_d^R|C_j)P(C_j), \tag{32}$$

Taking action a_i, by definition, will incur the loss $\lambda(a_i,C_j)$ when a particular $\mathbf{x}_d^R$ is observed, and its true class is C_j. The expected loss associated with an action a_i is called *Risk*, and $R(a_i|\mathbf{x}_d^R)$ is the conditional risk:

$$R(a_i|\mathbf{x}_d^R) = \sum_{j=1}^{J} \lambda(a_i|C_j)P(C_j|\mathbf{x}_d^R). \tag{33}$$

The risk related to any particular observation of $\mathbf{x}_d^R$ can be minimized by selecting an action a_i that minimizes the conditional risk. The *Bayes decision procedure* minimizes this risk. We call *decision rule* a function $a(\mathbf{x}_d^R)$ that associates an observed vector of measured features $\mathbf{x}_d^R$ to one of the C possible actions and assumes one of the values $a_1,..,a_C$. The overall risk R is the risk associated with a particular decision rule. It is expressed as:

$$R = \int R(a(\mathbf{x}_d^R)|\mathbf{x}_d^R)p_x(\mathbf{x}_d^R)d\mathbf{x}_d, \tag{34}$$

where $d\mathbf{x}_d$ is the notation for a d-space volume element. The risk is minimized if an action a_i that minimizes the conditional risk $R(a_i|\mathbf{x}_d^R)$ is selected. The resulting minimum risk is the best performance that can be achieved and is called *Bayes risk*. In classification problems, each state of nature is usually associated with a different class, and the action a_i, is usually interpreted as the decision that the true state of nature is C_i. If action a_i is taken and the true state of nature is C_j, then the decision is correct if $i = j$ and in error if $i \neq j$. If errors are to be avoided, it is natural to seek a decision rule that minimizes the probability of error.

In practice, the Bayes method requires at least knowledge about the prior probabilities and class-conditional PDFs to be employed. The PDFs of type $p_x(\mathbf{x}|C_j)$ can be either approximated numerically by deriving the normalized d-dimensional histogram of the sample data containing the vector of d characteristics being measured in each class or analytically by assuming that the PDFs of the d measured characteristics can be approximated by a d-dimensional distribution expressed by closed form equation, e.g., the normal distribution. Thus, the classification is made by considering the correct class, the one with the highest probability of being "true" with the lowest associated cost.

Despite its simplicity, the Bayesian method can achieve high accuracy levels when combined with kernel density estimation in statistical learning frameworks. Other works have proposed the Bayesian method to solve problems such as channel detection in smart-grid (TAO *et al.*, 2021), topology optimization for wireless sensor networks in complex terrains (OROZA *et al.*, 2021), backlogged devices number estimation in low-power wide-area networks (SEO *et al.*, 2021), and others.

Many works have proposed SI tools to solve problems in B5G. For instance, (LIU *et al.*, 2019) introduces a Sparse-Bayesian Inference (SBI) method to localize users in massive MIMO systems which is capable of exploiting the sparse and high-resolution nature of Angle-of-Arrival (AoA) and any available Statistical location information (SLI), to enhance the user localization accuracy. In (THOOTA; MURTHY, 2019), the *quantized variational Bayesian soft-symbol decoder* is presented. This algorithm aims to obtain the posterior beliefs of the transmitted bits in an uplink coded massive MIMO. This method is based on the variational Bayesian approach (BISHOP, 2006) in which the exact posterior distribution is approximated using a factorizable posterior distribution. Other works in literature, such as (CHERGUI *et al.*, 2022) , (REZAZADEH *et al.*, 2021) also utilize SI tools in B5G.

3 A RANDOM ACCESS PROTOCOL FOR CROWDED MASSIVE MIMO SYSTEMS BASED ON A BAYESIAN CLASSIFIER

In this chapter, the SUCRe enhanced by the Bayesian Classifier (SUCRe-BC) protocol is introduced. Although the SUCRe protocol is highly effective, it only considers the strongest user competing for a particular pilot to be the user whose signal strength is greater than half the sum of all users' signal strengths competing for the same pilot. As the number of users increases, the likelihood of a particular user being the strongest, with less than 50% of the sum of users' signal strengths competing for the same pilot, also increases. Therefore, the performance of the SUCRe protocol tends to be less effective as more false-negative cases occur. The SUCRe-BC protocol aims to reduce the number of false-negative cases by replacing the decentralized decision-making step (3rd step) of the SUCRe protocol with a new decentralized procedure using a BC.In the sections that follow, the system model will be described. The proposed approach is based on the Bayesian classifier, described in Section 3.2. To corroborate the effectiveness of our proposed method, the numerical results are comprehensively explored in Section 3.3.

The results demonstrate the superiority of the Bayesian method in terms of Average Number of Access Attempts (ANAA) and Fraction of Failed Access Attempts (FFAA) compared to the SUCRe protocol. In addition, the proposed method shows robustness regarding variation in the number of antennas or variation in the SNR level at the cell's border. Thus, making the SUCRe-BC a promising approach for replacing the SUCRe protocol.

3.1 SYSTEM MODEL

In this section, the system model of the proposed procedure is described. It is assumed that a BS with M antennas, located in the center of a hexagonal cell and operating in TDD mode, serves a set of User Equipments (UEs) with time and frequency resources split into coherence blocks of τ_c channel uses. The coherence blocks are divided into two categories: payload data blocks and RA blocks. The payload data blocks are used for UL and DL data transmission to the UEs in the set of active users, $\mathcal{A}_i$, that have been granted access to the BS exclusive resources. The RA blocks is dedicated for RA from inactive UEs (i.e., some of those in $\mathcal{U}i \backslash \mathcal{A}_i$) that wish to be granted access to the payload data blocks; that is, to be allocated a temporary dedicated pilot (BJÖRNSON *et al.*, 2017a).

We represent by $\mathcal{U}_i$, the set of all UEs inside cell i and $\mathcal{A}_i \subset \mathcal{U}_i$ the subset of $\mathcal{U}_i$

representing the active UEs. In a typical scenario where UEs are overcrowded, we have $|\mathcal{U}_i| \gg \tau_c$. Nonetheless, inactive UEs try to become active with a probability $P_a \leq 1$. Hence, we can also consider a scenario where $|\mathcal{A}_i| < \tau_c$ in which the BS can temporarily make orthogonal Payload data pilot (PDP) signals available to all active UEs during payload data transmission, by employing a GB RA protocol.

For this model, we focus on a central cell arbitrarily chosen and called cell 0, in which $\mathcal{K}_0 = \mathcal{U}_0 \backslash \mathcal{A}_0$ is the set of inactive UEs with cardinality $K_0 = |\mathcal{K}_0|$. It is also considered that the BS provides a quantity of τ_p orthogonal RA pilot signals $\{\psi_1, \psi_2, ...\psi_{\tau_p}\} \in \mathbb{C}^{\tau_p}$ for K_0 UEs to share, satisfying $\|\psi_t\|^2 = \tau_p$, $t \in \{1,2,...,\tau_p\}$. In each RA block, UEs that want to become active randomly choose one of the τ_p available RA pilot signals and make an access attempt by transmitting $\psi_{c(k)}$ with power $\rho_k > 0$, with $c(k) \in \{1,2,...,\tau_p\}$. The set $\mathcal{S}_t = \{k : c(k) = t, \rho_k > 0\}$ contains the indices of the UEs that transmit the pilot t, and thus $|\mathcal{S}_t|$ represents the number of UEs that choose the pilot ψ_t and follows a binomial distribution (BJÖRNSON *et al.*, 2017a):

$$|\mathcal{S}_t| \sim \mathcal{B}\left(K_0, \frac{P_a}{\tau_p}\right). \tag{35}$$

The condition of pilot t to be unused ($|\mathcal{S}_t| = 0$) is given by the probability $\left(1 - \frac{P_a}{\tau_p}\right)^{K_0}$. The probability of pilot t be selected by only one UE ($|\mathcal{S}_t| = 1$) with probability $K_0 \frac{P_a}{\tau_p}\left(1 - \frac{P_a}{\tau_p}\right)^{K_0 - 1}$. Therefore, an RA collision ($|\mathcal{S}_t| \geq 2$) at this arbitrary pilot occur with probability

$$1 - \left(1 - \frac{P_a}{\tau_p}\right)^{K_0} - K_0 \frac{P_a}{\tau_p}\left(1 - \frac{P_a}{\tau_p}\right)^{K_0 - 1}. \tag{36}$$

Before any UE can be allocated a temporary dedicated pilot, these collisions must be detected and resolved.

The vector $\mathbf{h}_k \in \mathbb{C}^M$ denotes the channel between UE $k \in \mathcal{K}_0$ and its BS. These channels are said to provide *channel hardening* as well as *asymptotic favorable* propagation conditions, which is mathematically expressed as:

$$\frac{\|\mathbf{h}_k\|^2}{M} \xrightarrow{M \to \infty} \beta_k, \forall k, \tag{37}$$

$$\frac{\mathbf{h}_k^{\mathrm{H}} \mathbf{h}_i}{M} \xrightarrow{M \to \infty} 0, \forall k,i, k \neq i \tag{38}$$

where β_k is a strictly positive value known by the k-th UE.

The SUCRe protocol has four steps that satisfy both (37) and (38) conditions. In the first step, the BS receives the signal $\mathbf{Y}_p \in \mathbb{C}^{M\times\tau_p}$ from the pilots sent by the UEs:

$$\mathbf{Y}_p = \sum_{k\in\mathcal{K}_0} \sqrt{\rho_k}\mathbf{h}_k\psi_{c(k)}^{\mathrm{T}} + \mathbf{B} + \mathbf{N}_p, \tag{39}$$

in which $\mathbf{B} \in \mathbb{C}^{M\times\tau_p}$ is the interference from adjacent cells, and $\mathbf{N}_p \in \mathbb{C}^{M\times\tau_p}$ is the noise matrix of the signal received by the BS with each element following a complex normal distribution, $\mathcal{CN}(0,\sigma^2)$. Then, BS correlates $\mathbf{Y}_p$ with ψ_t and obtains

$$\begin{aligned}\mathbf{y}_t &= \mathbf{Y}_p\tfrac{\psi_t^*}{\|\psi_t\|} = \textstyle\sum_{i\in\mathcal{S}_t}\sqrt{\rho_i}\|\psi_t\|\mathbf{h}_i + \mathbf{B}\tfrac{\psi_t^*}{\|\psi_t\|} + \mathbf{n}_t \\ &= \textstyle\sum_{i\in\mathcal{S}_t}\sqrt{\rho_i\tau_p}\mathbf{h}_i + \mathbf{B}\tfrac{\psi_t^*}{\|\psi_t\|} + \mathbf{n}_t,\end{aligned} \tag{40}$$

in which $\mathbf{n}_t = \mathbf{N}_p\frac{\psi_t^*}{\|\psi_t\|}$ is the effective noise with distribution $\mathcal{CN}(0,\sigma^2\mathbf{I}_M)$.

The ICI $\mathbf{B}$ term can be modelled as:

$$\mathbf{B} = \sum_l \mathbf{b}_l\mathbf{d}_l^{\mathrm{T}} + \sum_{t=1}^{\tau_p}\sum_{k\in\mathcal{S}_t^{\text{interf}}}\sqrt{\rho_{t,k}}\mathbf{g}_{t,k}\psi_t^{\mathrm{T}}, \tag{41}$$

where the first summation is over the interfering data transmissions carried out in neighboring cells that do not use the same set of pilots of cell 0. The l-th interferer has the channel $\mathbf{b}_l \in \mathcal{C}^M$ to the BS in cell 0 and transmits some random data sequence $\mathbf{d}_l \in \mathbb{C}^{\tau_p}$. The second summation is over the interferers in cells that reuse the pilots of cell 0, which also perform RA. The interferers that use pilot ψ_t are gathered in the set $\mathcal{S}_t^{\text{interf}}$, and member $k \in \mathcal{S}_t^{\text{interf}}$ has the channel $\mathbf{g}_{t,k}$ to the BS in cell 0 and uses the transmit power $\rho_{t,k}$.

$$\mathbf{B}\frac{\psi_t^*}{\|\psi_t\|} = \sum_l \mathbf{b}_l\frac{\mathbf{d}_l^{\mathrm{T}}\psi_t^*}{\|\psi_t\|} + \sum_{k\in\mathcal{S}_t^{\text{interf}}}\sqrt{\rho_{t,k}\tau_p}\mathbf{g}_{t,k}. \tag{42}$$

Assuming that all the interfering channels also satisfy the conditions in (37) and (38), denoted as $\frac{\|\mathbf{b}_l\|^2}{M} \to \beta_{w,l}$ and $\frac{\|\mathbf{g}_{t,k}\|^2}{M} \to \beta_{t,k}$, we obtain

$$\frac{|\mathbf{B}\frac{\psi_t^*}{\|\psi_t\|}|^2}{M} \xrightarrow{M\to\infty} \underbrace{\sum_l \beta_{\omega,l}\frac{|\mathbf{d}_l^{\mathrm{T}}\psi_t^*|^2}{\|\psi_t\|} + \sum_{k\in\mathcal{S}_t^{\text{interf}}}\rho_{t,k}\tau_p\beta_{t,k}}_{\omega_t}. \tag{43}$$

From Equations (40), (37) and (43) we derive:

$$\frac{\|\mathbf{y}_t\|}{M} \xrightarrow{M\to\infty} \sum_{i\in\mathcal{S}_t} \rho_i\beta_i\tau_p + \omega_t + \sigma^2. \tag{44}$$

In the second step of the SUCRe protocol, all UEs who sent pilot signals, is responded by the BS with a precoded signal $\mathbf{V} \in \mathbb{C}^{M\times\tau_p}$:

$$\mathbf{V} = \sqrt{q}\sum_{t=1}^{\tau_p} \frac{\mathbf{y}_t^*}{\|\mathbf{y}_t\|}\psi_t^{\mathrm{T}}, \tag{45}$$

in which q is the signal power available at this stage by the BS. The k-th UE then receives the signal $\mathbf{z}_k \in \mathbb{C}^{\tau_p}$:

$$\mathbf{z}_k^{\mathrm{T}} = \mathbf{h}_k^{\mathrm{T}}\mathbf{V} + \boldsymbol{\nu}_k^{\mathrm{T}} + \boldsymbol{\eta}_k^{\mathrm{T}}, \tag{46}$$

where $\boldsymbol{\nu}_k^{\mathrm{T}} \in \mathbb{C}^{\tau_p}$ is ICI and $\boldsymbol{\eta}_k^{\mathrm{T}}$ is the noise of the signal received by the UE in this step, with distribution $\mathcal{CN}(0,\sigma^2\mathbf{I}_{\tau_p})$. Then the UE correlates $\mathbf{z}_k$ with its chosen pilot ψ_t, resulting in

$$z_k = \mathbf{z}_k^{\mathrm{T}}\frac{\psi_t^*}{\|\psi_t\|} = \sqrt{q\tau_p}\mathbf{h}_k^{\mathrm{T}}\frac{\mathbf{y}_t^*}{\|\mathbf{y}_t\|} + \boldsymbol{\nu}_k^{\mathrm{T}}\frac{\psi_t^*}{\|\psi_t\|} + \eta_k, \tag{47}$$

where $\eta_k \sim \mathcal{CN}(0,\sigma^2)$. We derive

$$\frac{z_k}{\sqrt{M}} = \sqrt{q\tau_p}\frac{(\mathbf{h}_k\mathbf{y}_t)^*}{M}\frac{1}{\sqrt{\frac{1}{M}\|\mathbf{y}_t\|^2}} + \frac{\nu_k^{\mathrm{T}}\psi_t^*}{\sqrt{M}\|\psi_t\|} + \frac{\eta_k}{\sqrt{M}} \xrightarrow{M\to\infty} \frac{\sqrt{\rho_k q}\beta_k\tau_p}{\sqrt{\sum_{i\in\mathcal{S}_t}\rho_i\beta_i\tau_p + \omega_t} + \sigma^2}, \tag{48}$$

from the asymptotic favorable propagation, the convergence in (44), and the fact that noise does not increase with M. We also assume that ICI η_k is unaffected by M. Let α_t be the sum of the signal strengths and interference received by the BS during the first stage of the protocol for each pilot t, according to eq. (40). α_t can be expressed as:

$$\alpha_t = \sum_{i\in\mathcal{S}_t}\rho_i\beta_i\tau_p + \omega_t, \tag{49}$$

where ω_t is the interference derived in (43).

As proposed in (BJÖRNSON *et al.*, 2017a), the value of α_t can be estimated as:

$$\hat{\alpha}_{t,k} = \max\left(\left[\frac{\Gamma(M+\frac{1}{2})}{\Gamma(M)}\right]^2 \frac{q\rho_k\beta_k^2\tau_p^2}{[\Re(z_k)]^2} - \sigma^2, \rho_k\beta_k\tau_p\right), \tag{50}$$

where function $\max(\cdot,\cdot)$ outputs the maximum of two values and $\Gamma(\cdot)$ is the gamma function.

In the third step, each UE k determines whether or not to retransmit the pilot signal based on its own average channel gain β_k. The goal here is to allow only one competing UE (the strongest one) to retransmit in order to resolve the collision and connect this UE to the network. When $\mathcal{R}_k$ is true, the pilot is retransmitted; when $\mathcal{I}_k$ is true, the pilot is not retransmitting:

$$\mathcal{R}_k : \rho_k\beta_k\tau_p > \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{51}$$

$$\mathcal{I}_k : \rho_k\beta_k\tau_p \leq \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{52}$$

where $\epsilon_k \in \mathbb{R}$ is a bias parameter that can be used to adjust the system behavior; for instance, it can be employed for maximizing the average number of resolved collisions or to minimize the occurrence of false positives (or negatives). The bias term can also be used to ensure that at most one UE will transmit the pilot in the third step of the SUCRe protocol when $\epsilon_k > 0$ for all k (BJÖRNSON *et al.*, 2017a). Björnson *et al.* (2017a) proposes a suitable value for ϵ_k:

$$\epsilon_k = \frac{\delta\beta_k}{\sqrt{M}} + \frac{\bar{\omega}}{2}, \tag{53}$$

where δ is standard deviations of $\frac{\|\mathbf{h}_k\|^2}{M}$ centered around its mean value β_k and $\bar{\omega}$ is the average UL interference given by:

$$\bar{\omega} = \mathbb{E}\left\{ \frac{\|\mathbf{B}\frac{\psi_t^*}{\|\psi_t\|}\|^2}{M} \right\}, \tag{54}$$

where the expectation is computed with respect to user locations and shadow fading realizations are assumed to be known at the UE. The bias term is further explored in Pereira *et al.* (2021), where it is optimized and combined with a NOMA-RA protocol.

The fourth step of the SUCRe method comprises granting access to network resources to the UE that successfully retransmits its pilot after the collision has been resolved. All UEs that failed to be granted access to the BS resources try to connect again in the next RA block with a probability P_r until the maximum number of access attempts is reached. If the maximum number of access attempts is reached without the UE having gained access to the exclusive BS channel, the UE automatically gives up its attempts.

The soft-SUCRe (s-SUCRe) protocol of (MARINELLO; ABRÃO, 2019) follows a similar framework. However, instead of applying the hard decision retransmission criterion described in (51) and (52), each UE retransmits or not its chosen pilot in the third step according to its probability of being the strongest contender, which is derived in (MARINELLO; ABRÃO, 2019, eq. (16) and (17)).

3.2 BAYESIAN CLASSIFIER

The Bayesian method is a classical inference statistical tool that compares the weighted PDFs of numerous classes and the costs associated with them and then selects the class having the highest odd of being true with the lowest cost associated with it. In this section, we show how the Bayesian classifier can be used to resolve pilot collisions in crowded massive MIMO networks using the strongest user criterion. To begin, we rearrange inequations (51) and (52) as

follows to obtain an input variable for the proposed BC:

$$\mathcal{R}_k : \phi_k > 0.5, \tag{55}$$

$$\mathcal{I}_k : \phi_k \leq 0.5, \tag{56}$$

in which the $\phi_k \in [0,1]$ factor represents the proportion of the signal strength of the k-th UE among the contenders:

$$\phi_k = \frac{\rho_k \beta_k \tau_p - \epsilon_k}{\hat{\alpha}_{t,k}}. \tag{57}$$

Resolving the pilot collisions using the strongest UE retransmission rule can be regarded as a classification task to be carried out at the UEs' side. Thus, we can define two classes for the UEs: C_0 that represents the class of UEs that are not the strongest contenders for their chosen pilots, and C_1 which represents the class of the strongest UEs, comprehending the set $\mathcal{C} = \{C_0, C_1\}$. We also define the state of each UE k as $\Omega_k \in \mathcal{C}$. Based on the ϕ_k estimate obtained after step 2 as in (57), the proposed BC operates by seeking the class that maximizes the *a posteriori* probability of the k-th UE as:

$$\begin{aligned} \widehat{\Omega}_k &= \underset{C_\ell \in \mathcal{C}}{\arg\max} \quad P(C_\ell|\phi_k), \\ &= \underset{C_\ell \in \mathcal{C}}{\arg\max} \quad \frac{p_x(\phi_k|C_\ell)P(C_\ell)}{p_x(\phi_k)}, \\ &= \underset{C_\ell \in \mathcal{C}}{\arg\max} \quad p_x(\phi_k|C_\ell)P(C_\ell). \end{aligned} \tag{58}$$

The Bayesian method is then proposed to statistically map and classify the values of ϕ_k based on the possibility of the existence of UEs $\in C_1$ with $\phi_k \leq 0.5$, where C_1 is the class of the strongest UEs. Although it is possible to calculate the distribution of ϕ_k in (57), using Meijer's G-function, as the distribution of the ratio of two random variables, we opted for an empirical estimate of the distribution of ϕ_k for the following reasons. Firstly, the process of calculating the distribution of the ratio of two random variables using Meijer's G-function is complex and involves a multi-dimensional integral and the product of hypergeometric functions. Secondly, an empirical approach is a more flexible and robust method for estimating PDFs, as it is non-parametric and does not make any assumptions about the underlying distribution of the data. Finally, with a large enough number of data samples, in the order of millions, we can achieve a relatively precise estimation of the ϕ_k distribution. However, further works are encouraged to explore an analytic approach for calculating the PDFs related to the proposed BC.

We conduct an empirical statistical analysis with the objective of obtaining the joint PDF $p_x(\phi_k,C_\ell) = p_x(\phi_k|C_\ell)P(C_\ell)$, $\ell \in 0,1$ numerically. This means that we aim to find the joint probability density function of the variable ϕ_k and the class C_ℓ, which is given by the product of the conditional PDF $p_x(\phi_k|C_\ell)$ of ϕ_k given C_ℓ and the probability $P(C_\ell)$ of a UE belonging to class C_ℓ, where $\ell \in 0,1$. The BC training data is generated using the MATLAB 2020a software, following a simulation setup similar to the one proposed and publicly shared by the authors of the SUCRe protocol (BJÖRNSON *et al.*, 2017a). The numerical parameters for data collection in the SUCRe protocol simulation is shown in Table 3.

The SUCRe-BC simulation parameters are summarized in Table 4. The column "Value" displays the specific value or range of values used for each parameter, while the column "Description" provides an explanation of the purpose of the parameter, along with the indication of Figures that depict the results of simulations that employ the respective value or range of values of the parameter. Parameters without indication of Figures are applied in all simulations of the section.

Table 3 – Numerical Parameters for Data Collection in SUCRe Protocol Simulation

Parameter	Value	Description
M	100	Number of BS antennas in the center and neighboring cells
P_a	0.001	Transmission probability
P_r	0.5	Probability of trying again in the next RA block
τ_p	10	Number of available RA pilot sequences
ρ	27 dBm	Transmit power of UEs
q	27 dBm	Transmit power of the BS
σ^2	-98.65 dBm	Noise variance
δ	-1	Number of standard deviations in Equation (53)
K_{ici}	10	Number of active users in the neighboring cells
K_0	from 100 to 40000	Variation in the number of UEs in the center cell
Edge SNR	0 dB	Edge SNR in the center cell
	6	Number of neighboring cells
R	250 m	Radius of the center cell
R	250 m	Radius of the neighboring cells
	27 dBm	Transmit power of UEs in adjacent cells
	10 dB	Shadow fading standard deviation
	5G sub-6 GHz	Band of operation
	10000	Number of Monte-Carlo realizations
	10	Maximum number of connection attempts before the UE gives up

The offline training procedure we conducted consists in generating 10×10^6 examples of ϕ_k values labeled in terms of C_ℓ, with $\ell = 0, 1$, equally distributed between the different K_0 values. Then, we apply MATLAB function *histcounts* configured with 38-bin divisions and *probability* normalization on the training data. Such normalization implies that the area under graphs (a) and (b) of the Figures add up to one, since $\int_0^1 p_x(\phi_k,C_0)d\phi_k + \int_0^1 p_x(\phi_k,C_1)d\phi_k = 1$.

Table 4 – Simulation parameters of the proposed SUCRe-BC

Parameter	Value	Description
M	100	Number of BS antennas in the center and neighboring cells (Figures 11, 12, 13,14)
M	from 1 to 150	Variation in the number of BS antennas in the center and neighboring cells (Figure 15.a)
P_a	0.001	Transmission probability
P_r	0.5	Probability of trying again in the next RA block
τ_p	10	Number of available RA pilot sequences
ρ	27 dBm	Transmit power of UEs
q	27 dBm	Transmit power of the BS
σ^2	-98.65 dBm	Noise variance
δ	-1	Number of standard deviations in Equation (53)
K_{ici}	10	Number of active users in the neighboring cells
K_0	18000	Number users in the center cell (Figure 15)
K_0	from 100 to 30000	Variation in the number of UEs in the center cell (Figures 11, 12, 13,14)
Edge SNR	0 dB	Edge SNR in the center cell (Figures 11, 12, 13,14 and 15)
Edge SNR	from -10 to 10 dB	Variation of the edge SNR in the center cell (Figures 15.b)
	38	Number of bin divisions of the estimated $p_x(\phi_k\|C_\ell)$ and $p_x(\phi_k)$ histograms
	10×10^6	Number of ϕ_k labeled samples
	6	Number of neighboring cells
R	250 m	Radius of the center cell
R	250 m	Radius of the neighboring cells
	27 dBm	Transmit power of UEs in adjacent cells
	10 dB	Shadow fading standard deviation
	5G sub-6 GHz	Band of operation
	10	Maximum number of connection attempts before the UE gives up

Besides, the choice for a number of examples of 10×10^6 is empirical, trading off a training execution complexity of a few minutes while obtaining smooth PDF curves. The choice for a 38-bin divisions configuration is due to empirical tests with different bin divisions numbers ranging from 20 up to a 100, showing that a 38-bin configuration was the optimum choice by ensuring the greatest accuracy of the results. The procedure is initially performed without ICI, obtaining Figure 11, and then repeated with ICI, obtaining Figure 12.

Figure 11 shows the estimate of the PDFs of the ϕ_k values conditioned to the occurrences of each class, weighted by the class probabilities, *i.e.*, $p_x(\phi_k|C_\ell)P(C_\ell)$, $\ell = 0, 1$, for the scenario without ICI. The Figure also depicts the crossover point between both curves, pointing out the ϕ_k value at which $p_x(\phi_k|C_1)P(C_1)$ becomes greater than $p_x(\phi_k|C_0)P(C_0)$, which occurs for $\phi_k \approx 0.2942$.

Similarly, Figure 12 presents a similar result for the cases with ICI. In this scenario, the crossover point turns to occur for $\phi_k \approx 0.2853$.

The results in Figure 11 and Figure 12 were obtained by configuring the MATLAB function *histcounts* with 38-bin divisions and probability normalization weighted by the like-

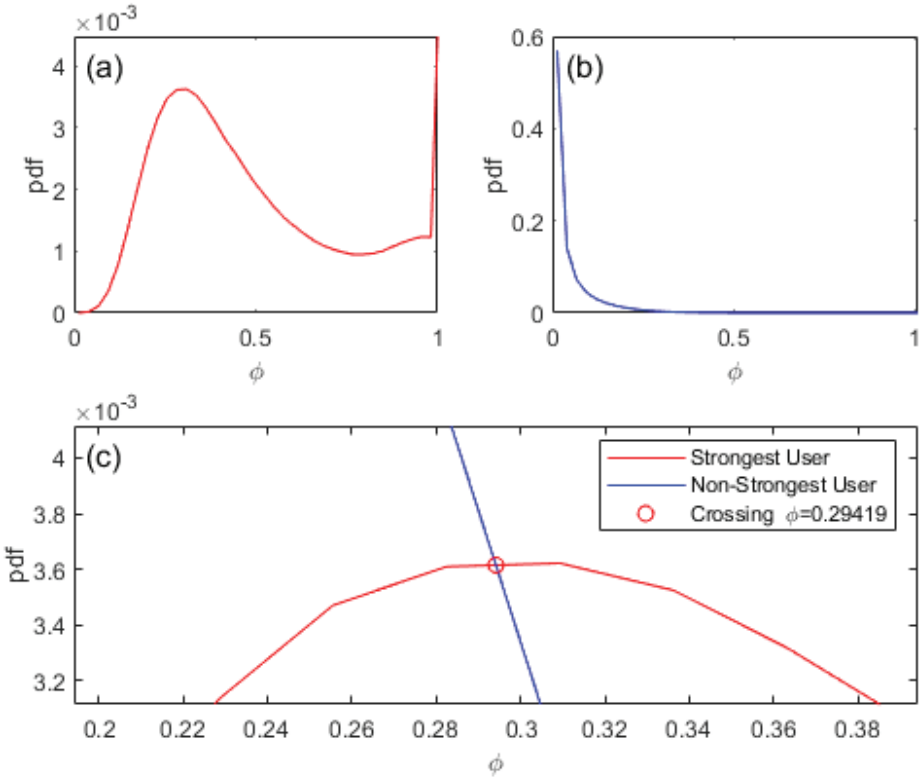


Figure 11 – $p_x(\phi_k|C_\ell)P(C_\ell)$**, (a)** $\ell = 1$**, (b)** $\ell = 0$**, and (c) the crossover point for the scenario without ICI. With** $K_{\mathbf{ici}} = 10$ $M = 100$ **and** K_0 **varying from** 100 **to** 30000**.**

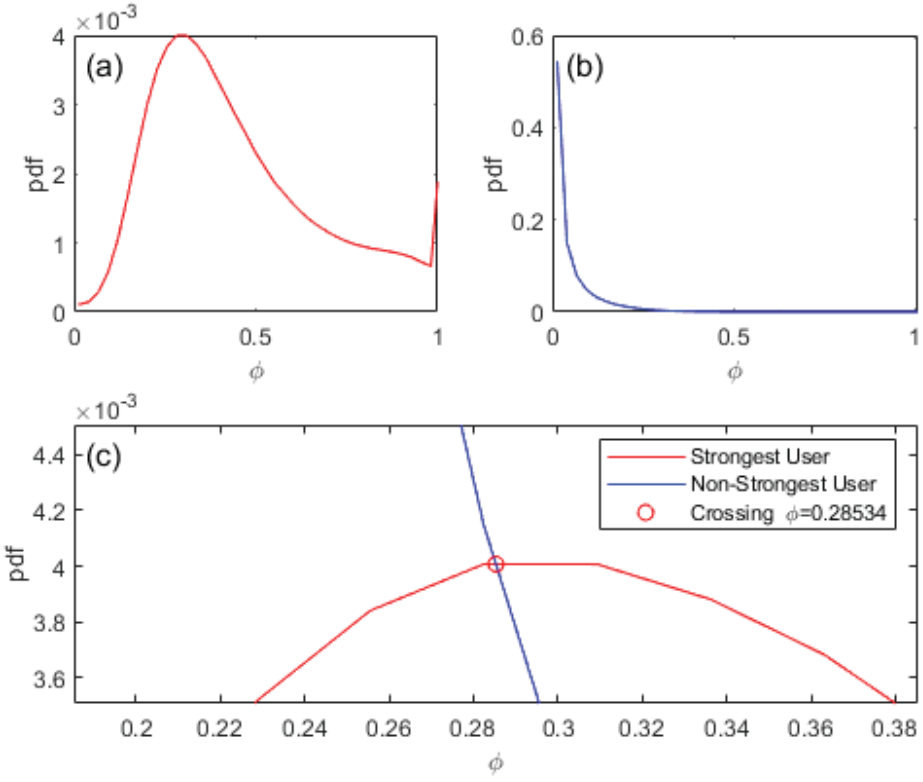


Figure 12 – $p_x(\phi_k|C_\ell)P(C_\ell)$**, (a)** $\ell = 1$**, (b)** $\ell = 0$**, and (c) the crossover point for the scenario with ICI. With** $K_{\mathbf{ici}} = 10$**,** $M = 100$ **and** K_0 **varying from** 100 **to** 30000**.**

lihood of their respective true class. As a result, the area under graphs (a) and (b) equals one. Interestingly, the above analysis shows that using BC as a criterion for UEs retransmitting their pilots in step 3 of the RA protocol is equivalent to changing the SUCRe protocol's retransmission rule decision threshold of 0.5 to a more refined value of 0.2942 in the absence of ICI, or 0.2853 in the presence of ICI. As a result, the inequations (55) and (56) can be replaced by

$$\begin{aligned} & \text{(w/o ICI)} && \text{(with ICI)} \\ \mathcal{R}_k: \quad & \phi_k > 0.2942, && \phi_k > 0.2853, \qquad (59) \\ \mathcal{I}_k: \quad & \phi_k \leq 0.2942, && \phi_k \leq 0.2853. \qquad (60) \end{aligned}$$

In other words, while the SUCRe protocol's goal is to resolve pilot collisions by allowing only the strongest UE to retransmit their pilots, it does so by evaluating $\phi_k > 0.5$. Indeed, $\phi_k > 0.5$ is a sufficient condition for a UE to be the strongest contender, but it is not necessary, leading to an increase in false-negatives and a deterioration in connectivity performance. On the other hand, as shown in the following section, the SUCRe-BC uses a more refined decision threshold near 0.29, which significantly improves RA performance.

3.3 NUMERICAL RESULTS

This section numerically evaluates the performance of the investigated RA protocols. First, we present and compare the RA performance of the protocols, and then we present and compare the classification accuracy of the schemes.

3.3.1 Confusion Matrix Interpretation

Before presenting the results, which are presented in the form of graphs and tables of confusion matrices, this section presents a Confusion Matrix (CM) model to facilitate the understanding of the results presented in this format. Table 5 presents a CM model to be interpreted as follows:

The column on the far right of the matrix presents the positive predictive values (PPVs), also known as precision, which are the percentages of all the samples predicted to belong to each state belonging to $\widehat{\Omega}_k$ that are correctly classified. In other words, it shows the classifier's accuracy for each output class $\widehat{\Omega}_k = C_0$ or $\widehat{\Omega}_k = C_1$. The same column also shows the false discovery rates (FDRs), which are the percentages of all the samples predicted to belong to each

Table 5 – Confusion Matrix Model

Predicted Class				
	C_0	**T_0**	**F_0**	**PPV_0 FDR_0**
	C_1	**F_1**	**T_1**	**PPV_1 FDR_1**
	total	**TPR_0 FNR_0**	**TPR_1 FNR_1**	**AC**
		C_0	C_1	**total**
		Actual Class		

class that are incorrectly classified:

- **PPV_0** - Indicates the PPV rate of all the samples predicted to belong to class C_0 .
- **PPV_1** - Indicates the PPV rate of all the samples predicted to belong to class C_1 .
- **FDR_0** - Indicates the FDR of all the samples predicted to belong to class C_0.
- **FDR_1** - Indicates the FDR of all the samples predicted to belong to class C_1.

The row at the bottom of the matrix shows the true positive rates (TPRs), also known as recall, which are the percentages of all the samples of each state belonging to $\widehat{\Omega}_k$ that are correctly classified. The row at the bottom also shows the false negative rates (FNRs), which are the percentages of samples belonging to each class that are incorrectly classified.

- **TPR_0** - Indicates the TPR of all the samples belonging to class C_0.
- **TPR_1** Indicates the TPR of all the samples belonging to class C_1.
- **FNR_0** - Indicates the FNR of all the samples belonging to class C_0.
- **FNR_1** Indicates the FNR of all the samples belonging to class C_1.
- **AC** - Indicates the overall percentage accuracy of the proposed classifier.

3.3.2 Classification Performance

Tables 6 and 7 evaluate the classification performance for the scenario without ICI in terms of confusion matrices. Table 6 shows the results when using the SUCRe protocol's

Table 6 – CM of SUCRe w/o ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	6829732 **93.1%**	293539 **4.0%**	**95.9%** **4.1%**
C_1	2977 **0.0%**	210311 **2.9%**	**98.6%** **1.4%**
total	**100.0%** **0.0%**	**41.7%** **58.3%**	**96.0%** **4.0%**

Table 7 – CM of the SUCRe-BC w/o ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	6763743 **92.2%**	127111 **1.7%**	**98.2%** **1.8%**
C_1	68966 **0.9%**	376739 **5.1%**	**84.5%** **15.5%**
total	**99.0%** **1.0%**	**74.8%** **25.2%**	**97.3%** **2.7%**

decision threshold of $\phi_k = 0.5$, whereas Table 7 shows the results when using the proposed BC approach, with the obtained new threshold of $\phi_k = 0.2942$.

The results in Table 6 show that the classification success rate of UE of class C_0 is 100% for the decision threshold of $\phi_k = 0.5$ used in the SUCRe protocol. Nevertheless, when $\phi_k \leq 0.5$, the successful classification rate among $C1$ UEs is 41.7%, indicating too many false negatives. Moreover, the precision of C_0 and C_1 classifications is 95.9% and 98.6%, respectively. The overall success rate of successful classifications is 96%.

Table 7 shows the outcomes when the decision threshold is $\phi_k = 0.2942$. The classification precision of C_0 and C_1 outputs is 98.2% and 84.5%, respectively. The percentages of correct predictions for class C_0 UEs are 99% and 74.8% for class C_1 UEs. The overall accuracy is 97.3%, which is higher than the results obtained for the $\phi_k = 0.5$ threshold.

The confusion matrices for the cases with ICI are shown in Tables 8 and 9. Table 8 displays the results when the decision threshold is set to $\phi_k = 0.5$. The results show an accuracy of 95.3% for output class C_0 and 99% for output class C_1. For UEs belonging to classes C_0 and C_1, the rate of successful classified states is 100% and 34.8%, respectively. The overall accuracy is 95.4%. When ICI is considered, the inferior classifier performance stands out.

For the SUCRe-BC with ICI, Table 9 shows the results when $\phi_k = 0.2853$ is used as the decision threshold. The results for SUCRe with ICI show an accuracy of 98.0% for output class C_0 and 83.9% for output class C_1. For UEs belonging to classes C_0 and C_1, the rate of successful classified states is 99.0% and 72.8%, respectively. The overall accuracy is 97.1%, which is slightly lower than the case without ICI.

In summary, the results in this subsection show that, when compared to the SUCRe protocol, the BC approach proposed in this work trades off a very slight increase in false-positive rates for a significant decrease in false-negative probabilities. Hence, using the BC increases the

Table 8 – CM of SUCRe with ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	6866246 **93.0%**	336464 **4.6%**	**95.3%** **4.7%**
C_1	1844 **0.0%**	179662 **2.4%**	**99.0%** **1.0%**
total	**100.0%** **0.0%**	**34.8%** **65.2%**	**95.4%** **4.6%**

Table 9 – CM of the SUCRe-BC with ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	6795980 **92.0%**	140362 **1.9%**	**98.0%** **2.0%**
C_1	72110 **1.0%**	375764 **5.1%**	**83.9%** **16.1%**
total	**99.0%** **1.0%**	**72.8%** **27.2%**	**97.1%** **2.9%**

correct classification for UEs of class C_1 significantly. The results point to 41.7% of SUCRe to 74.8% in the scenario without interference and 34.8% to 72.8% in the case with ICI. These results are also very beneficial to the RA performance, as shown in the following section.

3.3.3 Connectivity Performance

Here, we present the results of the SUCRe-BC in terms of ANAA and FFAA. Figures 13.a and 13.b show, respectively, the results of ANAA and FFAA for the cases without ICI and Figures 14.a and 14.b show, respectively, the results of ANAA and FFAA for the cases when ICI is considered. In both Figures, the result is presented in the following format: The black line represents the performance of the baseline protocol of (BJÖRNSON *et al.*, 2017a), which is an ALOHA-like protocol where pilot collisions are only handled by retransmission in later RA blocks. The red lines with "+" marker indicate the results obtained with the original SUCRe protocol of (BJÖRNSON *et al.*, 2017a), and the magenta lines with "□" marker indicate the results obtained with the s-SUCRe protocol of (MARINELLO; ABRÃO, 2019). The cyan lines with "○" marker indicate the results obtained with the ACBPC protocol of (MARINELLO *et al.*, 2020). The blue lines with "x" marker indicate the results obtained when employing the BC methodology proposed in this paper. The dotted line shows the results obtained for the cases with ICI. Both Figures show that the proposed BC method performs significantly better when the average number of accessing UEs exceeds the number of available pilots, which occurs when $K_0 > \tau_p/P_a = 10000$ UEs. The proposed protocol is also kept as simple as the original SUCRe protocol by avoiding any additional complexity or overhead. Furthermore, Figure 15 depicts the FFAA with variation in the number of BS, M, antennas and SNR at the cell border (edge SNR), defined as ρ/σ^2 and q/σ^2, with $\rho = q$. Although the proposed BC is trained with $M = 100$ and

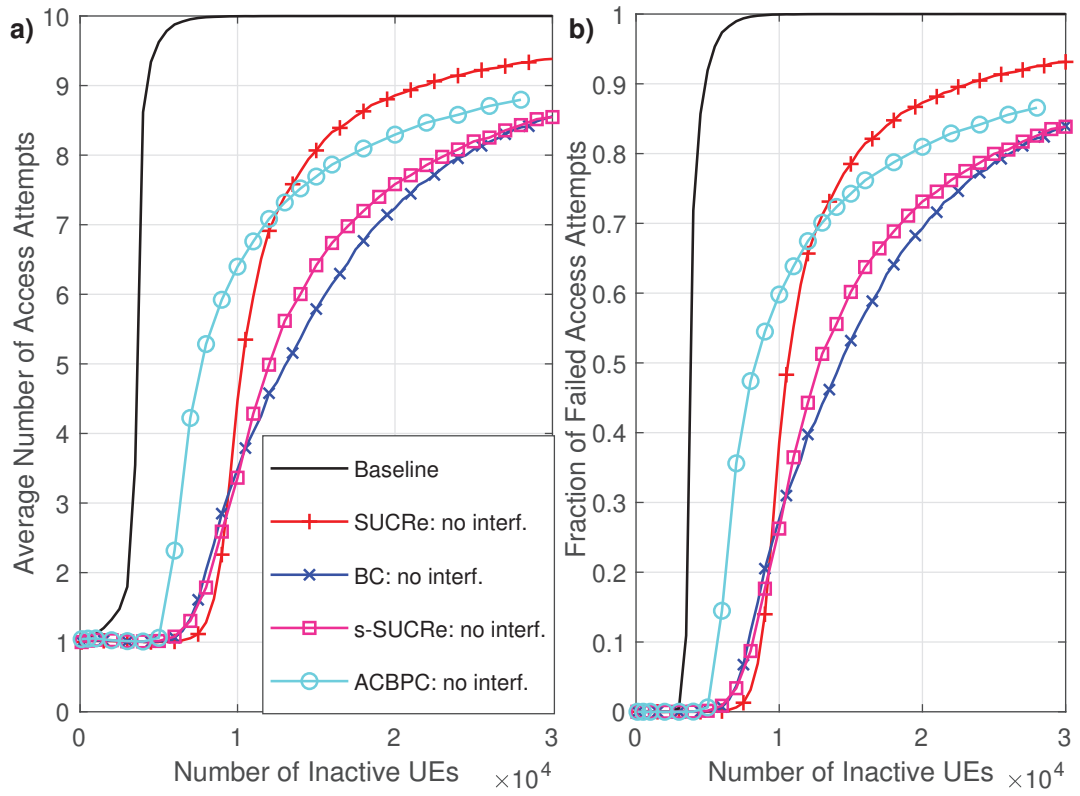


Figure 13 – Performance with K_0 variation: a) ANAA, and b) FFAA. w/o ICI for $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$, and 0 dB of edge SNR.

0 dB of SNR, the results show the robustness of the approach regarding the variation of such parameters.

3.4 CHAPTER CONCLUSIONS

This chapter proposes a statistical approach using a Bayesian classifier to optimize the SUCRe protocol. The classification results showed that our proposed methodology trades off a slight increase in false-positive rates for a significant decrease in false-negative probabilities, resulting in significantly higher correct classification probabilities for the strongest UEs. This behavior is very beneficial in terms of connectivity performance, which has been significantly improved by our proposed method over the SUCRe protocol in both scenarios with and without ICI. The most significant performance improvements have been found in the overcrowded scenario, with more expected accessing UEs than available RA pilots, *i.e.*, $K_0 > \tau_p/P_a$, indicating the proposed method as a promising approach in massive MIMO systems with very high UE density.

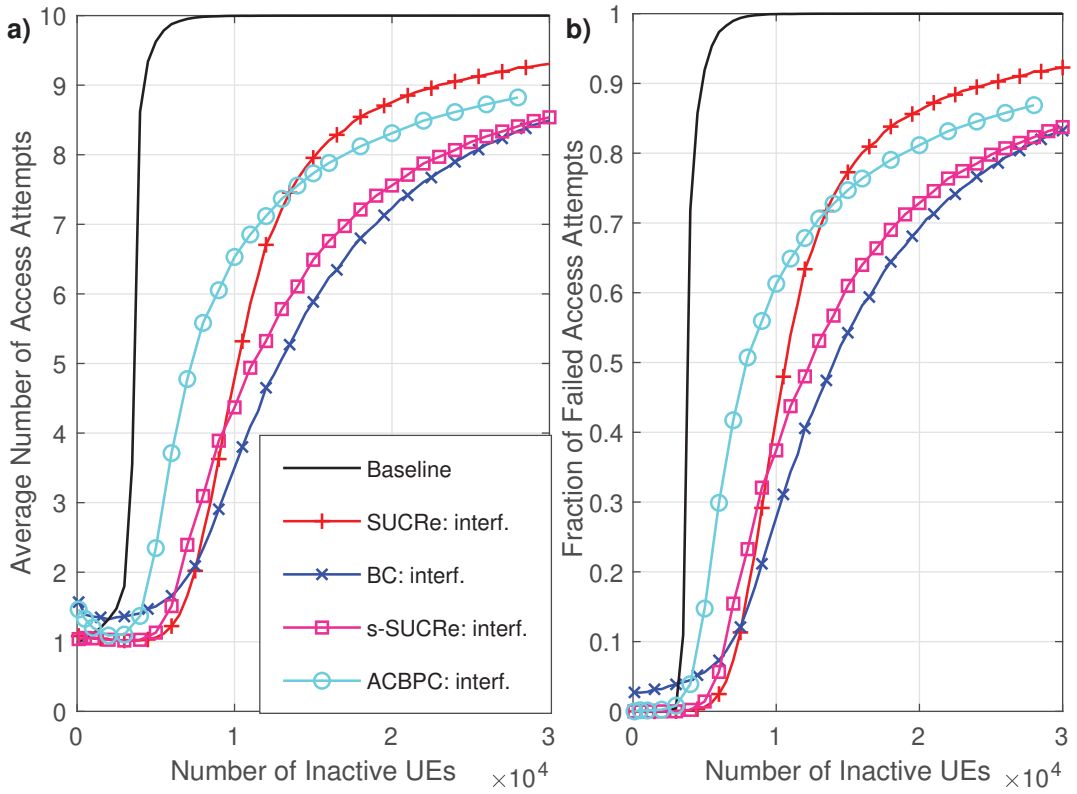


Figure 14 – Performance with K_0 variation: a) ANAA, and b) FFAA. With ICI for $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$, and 0 dB of edge SNR.

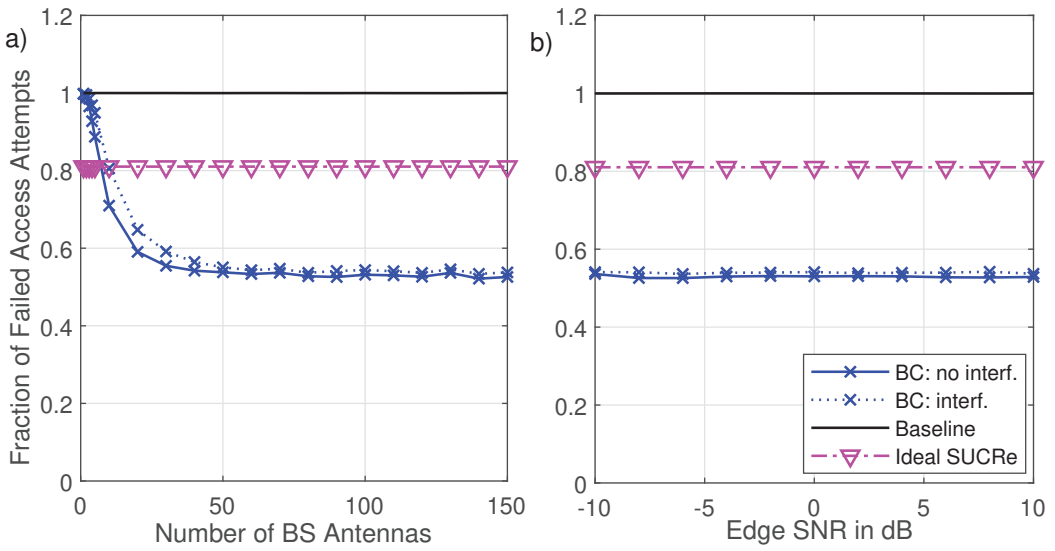


Figure 15 – FFAA *vs* a) Number of BS Antennas M, and b) Edge SNR . For $K_0 = 18000$, $\tau_p = 10$, $K_{\text{ici}} = 10$.

4 A NEURAL NETWORK-BASED RANDOM ACCESS PROTOCOL FOR CROWDED MASSIVE MIMO SYSTEMS

This chapter proposes an alternative for SUCRe-BC method introduced in chapter 3. The idea presented in this chapter is to employ an MLP NN-based method to conduct the same classification task and compare the results with the ones achieved by the SUCRe-BC approach and with other previously proposed methods. Similarly to the SUCRe-BC protocol, it aims to reduce the number of false-negative cases by replacing the decentralized decision-making step (3rd step) of the SUCRe protocol with a decentralized procedure. The main reasons for choosing an MLP NN are the following: **a)** The MLP NN is a simple method and is already well-validated by the literature. **b)** Once trained, a simple MLP NN can be embedded on low-cost hardware with relative ease. **c)** It is a method capable of producing better results than those achieved by the SUCRe-BC method. Although the Bayesian method can achieve great accuracy, the results depend on the considered classes' statistical distribution. On the other hand, the NN has the potential to predict with 100% accuracy, which is the class of a given variable, regardless of its statistical distribution.

The following sections describe the architecture, topology, and training process of the proposed NN. The validation results of the proposed NN are presented in terms of the number of hidden layer neurons and different learning rate values. The results of the proposed NN-based method indicate improved performance compared to the SUCRe-BC and similar protocols. Additionally, the robustness of the proposed method with respect to variations in the number of antennas and Edge SNR values is evaluated and demonstrates promising results.

4.1 THE SYSTEM MODEL AND THE MULTILAYER PERCEPTRON CLASSIFIER

In this chapter, the same scenario proposed in chapter 3 is investigated. Therefore, the same system model described in the section 3.1 is considered for the application of the MLP classifier.

4.2 NEURAL NETWORK CLASSIFIER

One of the most relevant features of artificial NNs is their capability to learn from the presentation of input data that expresses the system's behavior. Hence, after the network has

learned the relationship between inputs and outputs, it can generalize solutions, meaning that the network can produce an output close to the expected (or desired) output of any given input value. In this section, we show how an NN classifier can be applied to the random access problem in crowded massive MIMO networks under the strongest user criterion to resolve pilot collisions. Aiming to resolve pilot collisions under the strongest UE retransmission rule on the UEs' side, we define two classes for the UEs: C_0 represents the class of UEs that are not the strongest competitors for their chosen pilots, and C_1 represents the class of the strongest UEs, forming the set $\mathcal{C} = \{C_0, C_1\}$. We also define the state of each UE k as $\Omega_k \in \mathcal{C}$. Our method works by seeking an approximation $\widehat{\Omega}_k$ of the class that truly represents the one of the k-th UE. This can be accomplished through an MLP NN by estimating a function $\Omega_k = f(x_1,x_2)$ that maps the input values $x_1 = \rho_k \beta_k \tau_p$ and $x_2 = \hat{\alpha}_{t,k}$ to the state Ω_k of the k-th UE.

The steps to realize this method are the following: *A. Database Acquisition*, *B. Preprocessing*, *C. Neural Network Training* and *D. Validation*

4.2.1 Database Acquisition

The first step is to acquire the NNs training data. The database is generated from the simulation setup public shared by the authors of (BJÖRNSON *et al.*, 2017a), where the values of $\hat{\alpha}_{t,k}$, β_k and their respective transmission classifications are collected, according to:

$$d_k = \begin{cases} 1 & \text{if } \Omega_k = C_1 \\ 0 & \text{if } \Omega_k = C_0. \end{cases} \tag{61}$$

In this work, it was collected near 5×10^6 labeled training data from the simulation.

4.2.2 Preprocessing

In the preprocessing step, the data is prepared for being used as input for the NN. The preprocessing step, which includes data shuffling and normalization, is essential for ensuring the proper functioning of a NN. The data shuffling prevents overfitting, and normalization is important to ensure all input data have the same scale. Normalization is also important because many activation functions, such as sigmoid and hyperbolic tangent ($\tanh$), are sensitive to the scale of the input data. Normalization helps to ensure that the input data falls within the range of these functions.

Given the skewed [1] nature of the collected dataset (primary dataset), the input values are randomly shuffled using the MATLAB function *randperm* to ensure that a subset of the primary dataset with a sufficiently large number of elements will have members of both classes in a proportion near or equal to the one of the primary dataset. Next, 10×10^5 data samples are taken from the shuffled dataset, from which 80% are separated for the training step and 20% for the validation. Then, the input values x_1 and x_2 of both training and validation data are normalized according to:

$$\bar{x}_1 = \frac{\ln(x_1) - \min(\ln(x_1))}{\max(\ln(x_1)) - \min(\ln(x_1))}, \tag{62}$$

and,

$$\bar{x}_2 = \frac{\ln(x_2) - \min(\ln(x_2))}{\max(\ln(x_2)) - \min(\ln(x_2))}. \tag{63}$$

The normalization procedure mathematically described in Equations (62) and (63) are divided into two steps. First, the natural logarithm of the input data is taken to eliminate large numerical discrepancies (above six orders of magnitude) among the input data, improving the data resolution. Next, the resulting values are normalized to fit the closed interval $[0,1]$, matching the output range of the activation function in (71). Finally, both input values are grouped into a vector $\bar{\mathbf{x}}_k$, where a bias term b is also appended:

$$\bar{\mathbf{x}}_k = [b \quad \bar{x}_1 \quad \bar{x}_2]^T, \tag{64}$$

where $(\cdot)^T$ is the transpose operation.

4.2.3 Neural Network Training

The training process of an NN consists of applying the required ordinated steps for tuning its neurons' synaptic weights and thresholds in order to generalize the solutions produced by its outputs. In the proposed method, the normalized data, $\bar{\mathbf{x}}_k$, and the desired output value d_k associated with each $\bar{\mathbf{x}}_k$ sample are used as input data for training a fully connected MLP NN, in which each neuron in one layer is connected to every neuron in the subsequent layer. In the proposed method the NN have one hidden layer, as illustrated in Figure 16. The MLP NN

[1] The data skewed nature comes from the fact that the dataset has a number of data items labeled with C_0 much larger than the ones labeled with C_1.

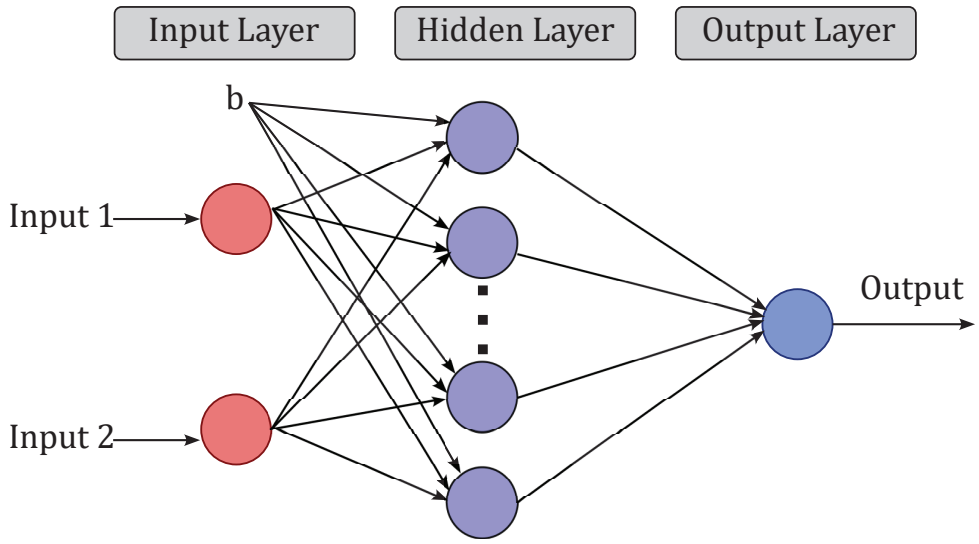


Figure 16 – MLP NN with one hidden layer

consists of a set of linear combinators, called neurons, that control scalar product operations between input vectors and sometimes a bias term b with synaptic weights to generate a result by applying a given activation function. The bias term b accounts as an additional input term for the hidden layer's neurons.

4.2.3.1 Backpropagation algorithm

The training process for a neural network is performed using the well-known backpropagation algorithm. This algorithm updates two weight matrices that connect the various layers of the network. The first weight matrix, $\mathbf{U_1} \in \mathbb{R}^{(L_I+1)\times L_H}$, connects the input layer with L_I neurons to the hidden layer with L_H neurons. The second weight matrix, $\mathbf{U_2} \in \mathbb{R}^{L_H \times L_O}$, connects the hidden layer with L_H neurons to the output layer with L_O neurons.

The backpropagation algorithm works as follows: first, the synaptic weights are randomly initialized, then two training steps alternate until the stopping criterion is reached. In the first step, the input signals are propagated toward the network's output layer. This is also known as the forward phase of training. The previous phase's output is compared to the desired output in the second step, and the weight matrices are adjusted backward. The weights of the output layer are adjusted first, then the weights of the previous layer are adjusted based on the weights of the output layer, and so on until the weights of the input layer are adjusted. This is also known as the backward phase. As the network in question has only two matrices of synaptic weights, the adjustment occurs only in the output layer and in the input layer. Equations (65) and (66) represent the adjustment of the weights in the output layer, and Equations (67) and (68) represent

the adjustment of the weights in the input layer (SILVA *et al.*, 2016). Where κ is the learning rate, $grad_j^{(2)}$ is the local gradient with respect to the j-th neuron of the second layer, $grad_j^{(1)}$ is the local gradient with respect to the j-th neuron of the first layer, d_j is the j-th desired output value, $o_j^{(2)}$ is the value of the j-th output provided by the network, $sig'(\cdot)$ is the derivative of the sigmoid function, $u_{ji}^{(2)}(t)$ is the synaptic weight between the j-th neuron of the output layer and the i-th neuron of the immediately previous layer, and $u_{ji}^{(2)}(t+1)$ is the value of the same synaptic weight updated by the generalized delta rule and in the same way $u_{ji}^{(1)}(t)$ and $u_{ji}^{(1)}(t+1)$ are respectively the previous and updated synaptic weights connecting the j-th neuron of the hidden layer and the i-th input value, $p_j^{(2)} = \sum_{i=1}^{L_H} u_{ji}^{(2)} sig(p_j^{(1)})$ is the activation potential of the j-th neuron of the second layer, $p_j^{(1)} = \sum_{i=1}^{L_I} u_{ji}^{(1)} x_i$ is the activation potential of the j-th neuron of the first layer, and x_i is the i-th input value.

$$grad_j^{(2)} = (d_j - o_j^{(2)}) \cdot sig'(p_j^{(2)}) \tag{65}$$

$$u_{ji}^{(2)}(t+1) = u_{ji}^{(2)}(t) + \kappa \cdot grad_j^{(2)} \cdot o_j^{(2)} \tag{66}$$

$$grad_j^{(1)} = -\sum_{k=1}^{L_O} (d_k^{(2)} - u_{ji}^{(2)}) \cdot sig'(p_j^{(1)}) \tag{67}$$

$$u_{ji}^{(1)}(t+1) = u_{ji}^{(1)}(t) + \kappa \cdot grad_j^{(1)} \cdot x_i \tag{68}$$

Given a learning rate κ, the backpropagation algorithm, proceeds iteratively by minimizing the mean square error (MSE) function between the desired outputs $\mathbf{o}_k$ and the actual output $\hat{\mathbf{o}}_k$ at each i-th training epoch. The choice for the MSE function follows the MLP NN model proposed by Silva *et al.* (2016, p. 41) for classification tasks:

$$\text{MSE}_i = \frac{1}{2P} \sum_{k=1}^{P} (\mathbf{o}_k - \hat{\mathbf{o}}_k)^2 , \tag{69}$$

where MSE_i is the MSE value at the i-th training epoch, and P is the total number of training samples. The training is considered complete when a given precision value, $\xi = \text{MSE}_i - \text{MSE}_{i-1}$ is achieved. Once trained, the MLP NN can be used to estimate the function $g(\cdot)$ as:

$$\hat{\mathbf{o}}_k = sig(\mathbf{U_2}^T \cdot sig(\mathbf{U_1}^T \bar{\mathbf{x}}_k)), \tag{70}$$

where $sig(\cdot)$ is the activation sigmoid function:

$$sig(x) = \frac{1}{1 + e^{-x}}. \tag{71}$$

Finally, the output $\hat{\mathbf{y}}_\mathbf{k}$ is associated to one of the output $\mathcal{C}$ classes. Thus, deriving the estimator $\widehat{\Omega}_k$:

$$\widehat{\Omega}_k = \begin{cases} C_0 & \text{if } \hat{\mathbf{o}}_k \leq 0.5 \ , \\ C_1 & \text{if } \hat{\mathbf{o}}_k > 0.5 \ . \end{cases} \tag{72}$$

4.2.4 Validation

Validation is a crucial step in demonstrating the ability of an NN to generalize its results over data that was not used during the training process. The aim of the validation in this work is to support the choice of the architecture, topology, and hyperparameters of the proposed NN binary classifier. To evaluate the performance of the proposed NN, the performance metrics *Recall*, *Precision*, *F-Measure*, and *Accuracy* are calculated using the set of data reserved for testing.

The *Recall* metric measures the proportion of actual instances of C_1 that were correctly classified as C_1. *Precision* indicates the ratio of C_1 predictions that were actually C_1 to the total number of C_1 predictions. The *F-Measure* is the harmonic mean of precision and recall, calculated as $2 \times (\textit{Precision} \times \textit{Recall})/(\textit{Precision} + \textit{Recall})$. Finally, *Accuracy* represents the proportion of correct predictions made by the MLP NN over all predictions. The results in both Tables are based on the classification of data generated from a simulation scenario with ICI.

The results of the evaluation are presented in Tables 10 and 11. These Tables show the performance of the NN binary classifier for different numbers of neurons in the hidden layer L_H and different learning rates κ.

Table 10 – Number of neurons test.

L_H	**Recall**	**Precision**	**F-Measure**	**Accuracy**
3	0.7414	0.8380	0.7867	0.9719
4	0.7414	0.8380	0.7867	0.9718
5	0.7485	0.8355	0.7896	0.9721
6	0.7567	0.8192	0.7867	0.9713
7	0.7079	0.8532	0.7738	0.9713
8	0.7941	0.7925	0.7933	0.9713
9	0.7763	0.8079	0.7918	0.9718
10	0.7689	0.8081	0.7918	0.9713

In Table 10 the number of neurons in the hidden layer L_H increases from 3 to 10 while the learning rate κ is fixed as 0.2 and the precision ξ is set to 10^{-7}. The bias term is set to $b = -1$, as recommended in (SILVA *et al.*, 2016). It is noteworthy from Table 10 that there is not a significant change of performance when varying L_H from 3 to 10. Also, there is no indication

of an increase in the overall accuracy metric. Nonetheless, there is a slight performance increase in the recall and the F-measure metrics. The F-measure is an important metric of performance for skewed data classification since it indicates that precision and recall metrics are balanced out.

The learning rate in Table 11 ranges from 0.01 to 0.2, with the hyperparameter L_H fixed at 5, ξ at 10^{-7} and $b = -1$. As κ increases from 0.01 to 0.2, the recall drops from 0.7738 to 0.7485, but precision improves slightly from 0.8113 to 0.8355. Despite these changes, accuracy and F-measure do not present significant changes.

Table 11 – Learning rate test.

κ	**Recall**	**Precision**	**F-Measure**	**Accuracy**
0.01	0.7738	0.8113	0.7921	0.9718
0.05	0.7556	0.8225	0.7876	0.9716
0.1	0.7280	0.8567	0.7871	0.9720
0.15	0.7617	0.8355	0.7880	0.9714
0.2	0.7485	0.8355	0.7896	0.9721

Figure 17 shows the convergence of the MSE function to its minimum value with training parameters set $H_L = 5$, $\kappa = 0.2$, $\xi = 10^{-7}$ and $b = -1$. The training is set up to stop when one of the two conditions occurs. The first stopping criterion is reaching the required precision value of $\xi = 10^{-7}$, and the second one is completing a predetermined number of training epochs, which is set to 1000. In Figure 17, the maximum number of epochs is achieved first, which stops the training at 1000 epochs of training. The MSE value achieved is 0.010263 with a precision below $\xi = 10^{-6}$.

In the following section, the third step of the SUCRe protocol is replaced with a MLP NN binary classifier with $L_O = 1$ neuron in the output layer, $L_I = 2$ neurons in the input layer, and $L_H = 5$ neurons in the hidden layer. The NN is trained with the backpropagation algorithm described in Subsection 4.2.3 with learning rate $\kappa = 0.2$. Two main training runs are carried out, the first with data collected in a scenario with ICI and the second with data collected in a scenario without ICI, yielding as the final result of the training process 4 weight matrices, 2 of them for the NN trained in the scenario with ICI and two of them for the NN trained in the scenario without ICI.

The L_H and κ values are chosen empirically among their tested values since their performance metrics do not show a significant difference. The number L_H is also kept relatively low as 5 because adding more neurons to the NN introduces more complexity making the algorithm computationally costly. For the same reason, a second hidden layer is not introduced. The bias term, b, is set to -1, as suggested in (SILVA *et al.*, 2016).

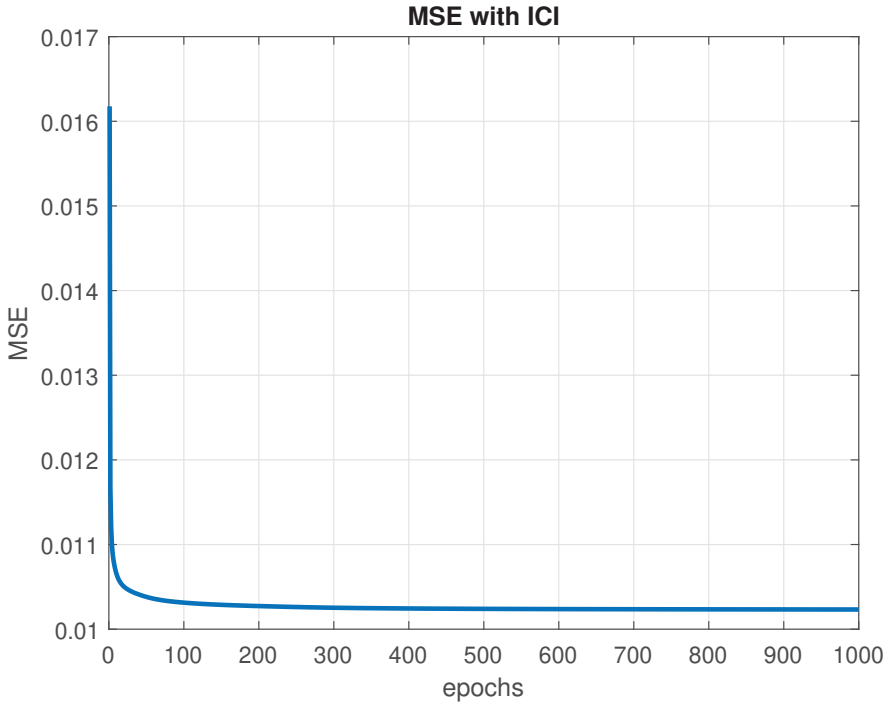


Figure 17 – MSE convergence. With ICI

4.3 NUMERICAL RESULTS

In this section, we present the results obtained with the proposed method in terms of confusion matrices, ANAA, and FFAA. The systems parameters used for collecting the training data of the proposed NN are summarized in Table 3. Table 12 shows the parameters of the proposed MLP NN based protocol. Similar to Table 4, the column "Value" displays the specific value or range of values used for each parameter, while the column "Description" informs the purpose of the parameter and indicate Figures that depict the results of simulations that employ the respective value or range of values of the parameter. Parameters without indication of Figures are applied in all simulations of the section.

We consider a system operating under the 5G sub-6 GHz band. The center cell has a radius of $R = 250$ m and a number of UEs K_0 varying between 100 and 30000 in increments of 100. The neighboring cells have the same radius and a fixed number of 10 active UEs each, matching the number of pilots and the expected number of active UEs in the neighboring cells. First, we provide results for the classification accuracy of the schemes, and then we present and compare the RA performance of the protocols.

Table 12 – Simulation parameters of the proposed GB RA NN-based protocol

Parameter	Value	Description
M	100	Number of BS antennas in the center and neighboring cells (Figures 18, 19, 20, 22)
M	from 1 to 100	Variation in the number of BS antennas in the center and neighboring cells (Figure 21)
P_a	0.001	Transmission probability
P_r	0.5	Probability of trying again in the next RA block
τ_p	10	Number of available RA pilot sequences
ρ	27 dBm	Transmit power of UEs
q	27 dBm	Transmit power of the BS
σ^2	-98.65 dBm	Noise variance
δ	-1	Number of standard deviations in Equation (53)
K_{ici}	10	Number of active users in the neighboring cells
K_0	28000	Number users in the center cell (Figure 21, 22)
K_0	from 100 to 30000	Variation in the number of UEs in the center cell (Figures 18, 19, 20)
K_0	from 25000 to 40000	Variation in the number of UEs in the center cell (Figure 20)
Edge SNR	0 dB	Edge SNR in the center cell (Figures 18, 19, 20, 20,21)
Edge SNR	from -8 to +8 dB	Variation of the edge SNR in the center cell (Figures 22)
	6	Number of neighboring cells
R	250 m	Radius of the center cell
R	250 m	Radius of the neighboring cells
	27 dBm	Transmit power of UEs in adjacent cells
	10 dB	Shadow fading standard deviation
	5G sub-6 GHz	Band of operation
	10	Maximum number of connection attempts before the UE gives up

4.3.1 Classification Performance

Tables 13 and 14 show the classification performance of the proposed MLP NN-based protocol in terms of confusion matrices. Table 13 presents the results for the scenario without ICI and Table 14 results for the scenario with ICI. The bottom row shows the successful classification rates of each state, the far-right column of the matrix shows the classification precision, and the far-right square at the bottom shows the overall accuracy of the classifier.

Table 13 – NN classifier w/o ICI

Predicted Class \ **Actual Class**	C_0	C_1	**total**
C_0	**184560** 92.3%	**3226** 1.6%	98.3% 1.7%
C_1	**1729** 0.9%	**10485** 5.2%	85.8% 14.2%
total	99.1% 0.9%	76.5% 23.5%	97.5% 2.5%

The results presented in Table 13 indicate that for the proposed MLP NN-based classifier, in a scenario without ICI, the classification success rate of UEs of class C_0 is 99.1%, and the successful classification rate among C_1 UEs is 76.5%, which is a number higher than the ones achieved by the SUCRe protocol, 41.7% and the ones achieved by the BC method, which is 74.8% (BUENO *et al.*, 2022). Besides, the precision of C_0 and C_1 classifications are 98.3% and 85.8%, respectively. The overall accuracy of successful classifications is 97.5%, which is also higher than the values achieved by the SUCRe protocol and the SUCRe-BC method, which are respectively 96% and 97.3% in a scenario without ICI.

Table 14 – NN classifier with ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	**184068** 92.0%	**3599** 1.8%	98.1% 1.9%
C_1	**1965** 1%	**10368** 5.2%	84.1% 15.9%
total	98.9% 1.1%	74.2% 25.8%	97.2% 2.8%

Table 14 presents the results for a scenario with ICI. The classification precision of C_0 and C_1 outputs are 98.1% and 84.1%, respectively. The percentages of correct predictions are 98.9% for class C_0 and 74.2% for class C_1. The overall accuracy is 97.2%, higher than the results obtained with the SUCRe protocol and the SUCRe-BC method.

4.3.2 Connectivity Performance

Figures 18 and 19, respectively, show the results of the metrics ANAA and FFAA in the following format: **a)** The baseline protocol of (BJÖRNSON *et al.*, 2017a), which is an ALOHA-like protocol where pilot collisions are only handled by retransmission in later RA blocks, is shown in a black line. **b)** The red lines marked with "$\circ$" indicate the results of the original SUCRe protocol presented in (BJÖRNSON *et al.*, 2017a). **c)** The cyan lines marked with "$\triangleright$" indicate the results of the the ACBPC protocol of (MARINELLO *et al.*, 2020). **d)** In the green lines marked with "x" are indicated the results obtained with the softSUCRe protocol of (MARINELLO; ABRÃO, 2019). **e)** The magenta line with "$\diamond$" marker represents the results obtained with the BC method of (BUENO *et al.*, 2022). **e)** Finally, the blue lines with the "$\square$"

marker indicate the results obtained when employing the MLP NN-based methodology proposed in this paper. The dotted line shows the results obtained for the cases without ICI. It is noteworthy the superiority of the proposed method. Compared to the SUCRe-BC method, for example, the proposed MLP NN-based method achieves a better performance when $K_0 \geq 25000$ UEs in the cases with ICI, as highlighted in Figure 20, where K_0 varies from 25000 up to 40000.

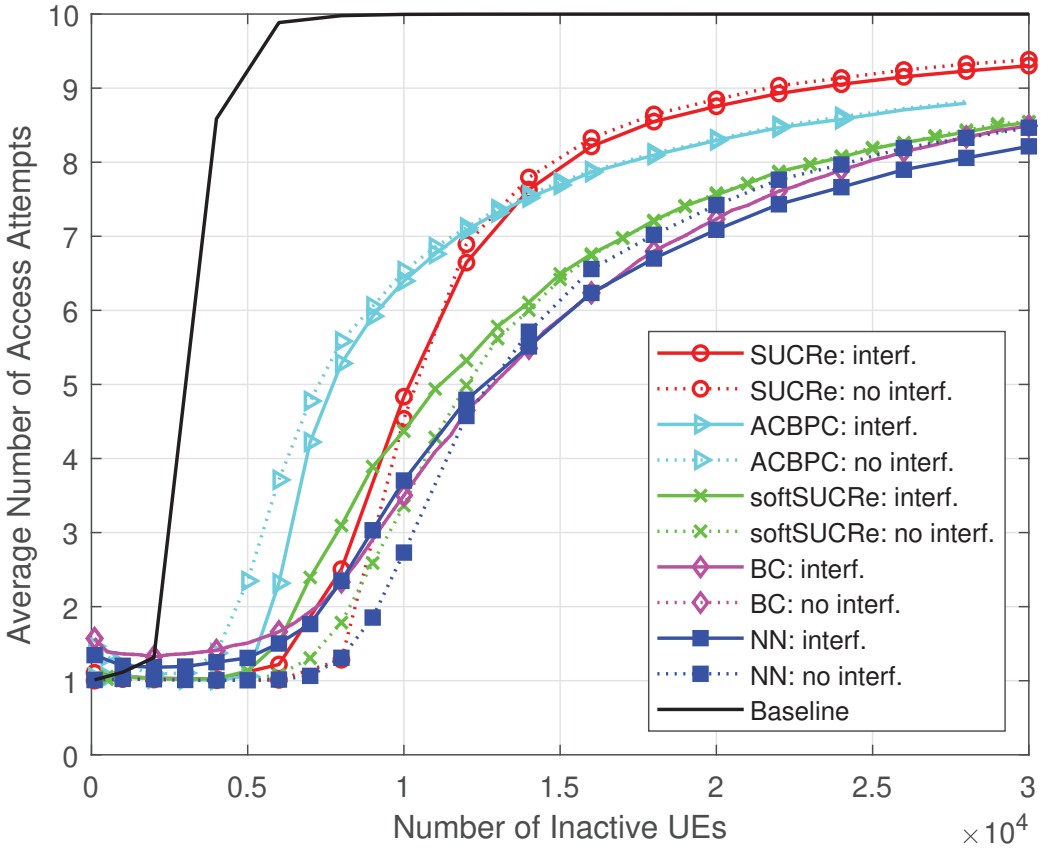


Figure 18 – ANAA $\times K_0$, for $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$, and 0 dB of edge SNR.

Figure 21 shows the results of the ANAA and FFAA when the number of antennas varies from $M = 1$ to $M = 100$ in steps of 2. Even though the NN-based method is trained with $M = 100$ antennas, the proposed method shows robustness under the variation in the number of BS antennas M, where a number of $M \approx 50$ antennas reveals to be sufficient to provide superior results than the SUCRe and BC methods in both with and without ICI scenarios.

In Figure 22 the edge SNR in dB, defined as ρ/σ^2 and q/σ^2, with $\rho = q$, varies from -8 dB to $+8$ dB. The ANAA and FFAA are taken for a fixed number $M = 100$ antennas and $K_0 = 28000$ UEs. In both charts the performance of the NN-based method in both scenarios with and without ICI, is superior than the performance of the SUCRe protocol, in the whole considered edge SNR range. Compared to the SUCRe-BC method, the NN-based method presents a superior

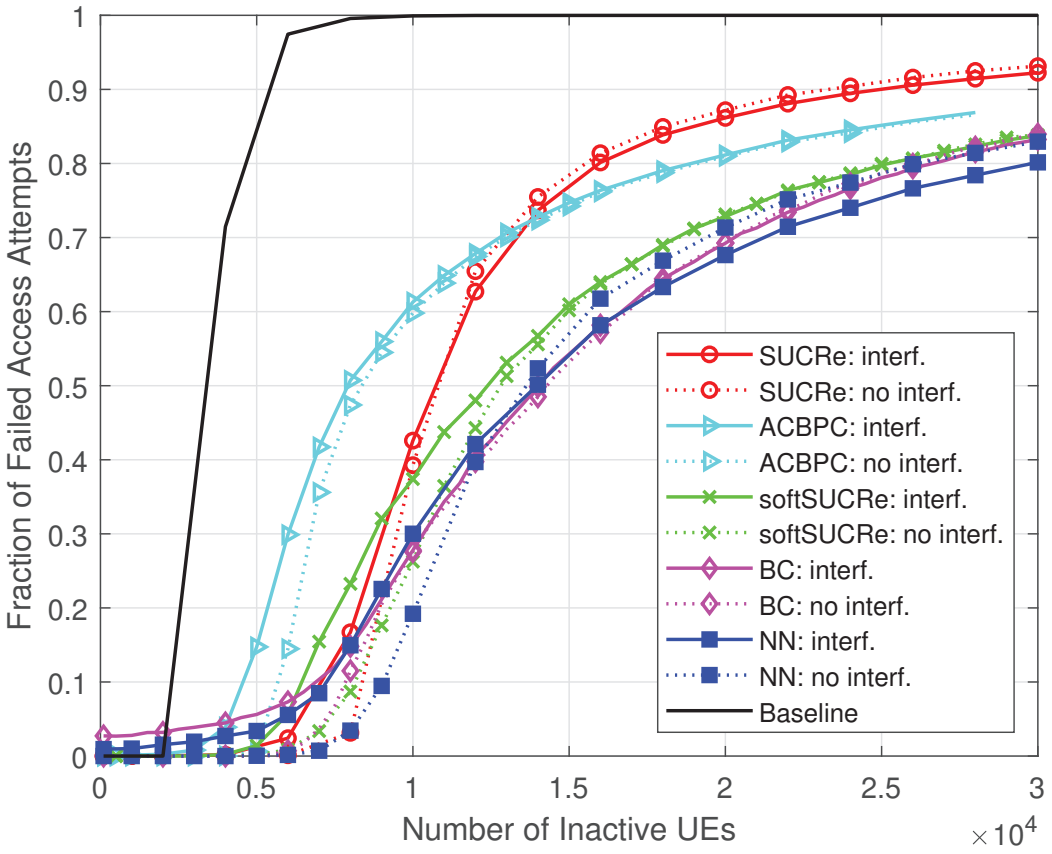


Figure 19 – FFAA $\times K_0$, for $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$, and 0 dB of edge SNR.

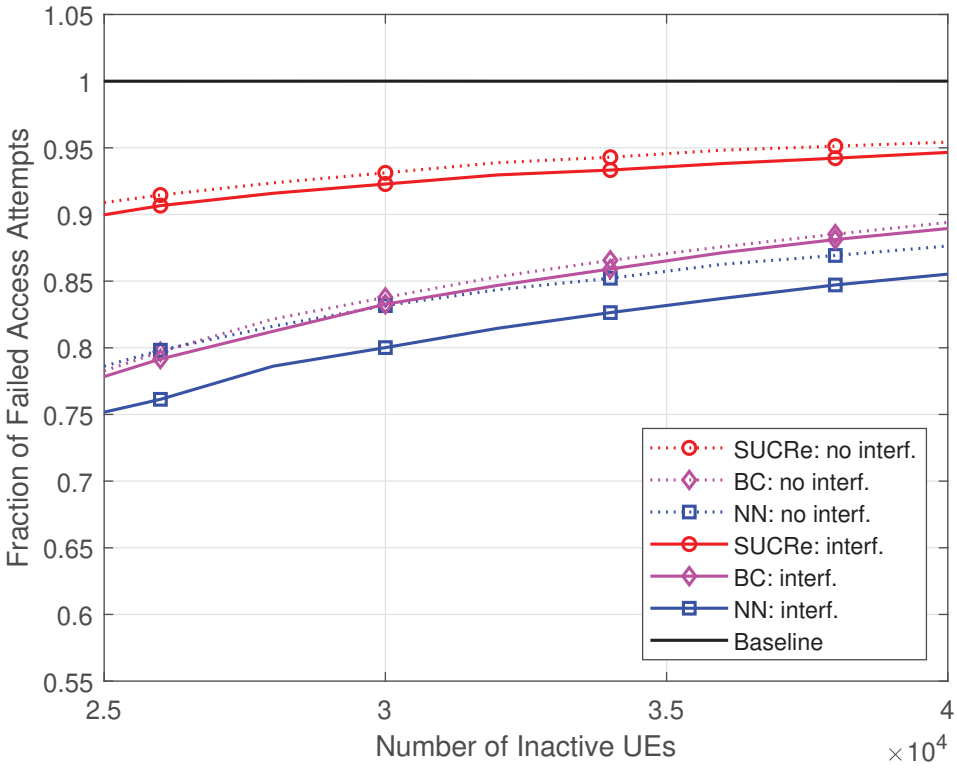


Figure 20 – FFAA $\times K_0$, for $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$, and 0 dB of edge SNR.

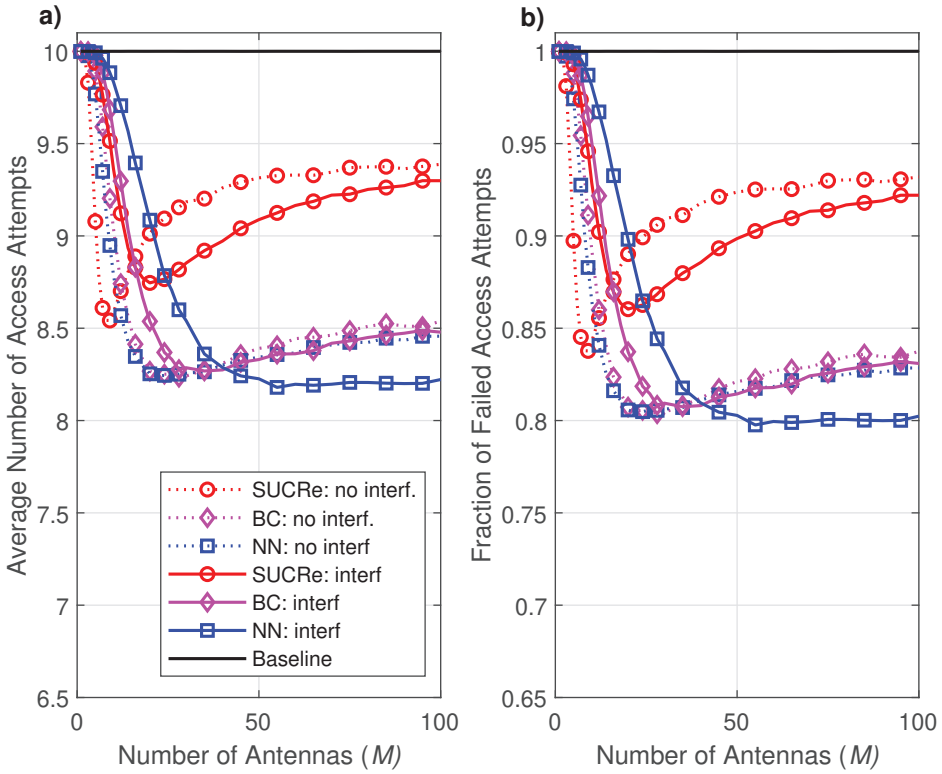


Figure 21 – Performance with M variation: a) ANAA, and b) FFAA. With $\tau_p = 10$, $K_{\text{ici}} = 10$, $K_0 = 28000$, and 0 dB of edge SNR.

performance in a scenario with ICI from ≈ -6 dB to $\approx +6$ dB, and in the cases without ICI, from the 0 dB to $\approx 6dB$.

4.4 CHAPTER CONCLUSIONS

In this chapter, it was implemented an NN-based method to optimize the SUCRe and other similar protocols. Based on our findings, we conclude that the proposed NN-based method outperforms the SUCRe protocol in both scenarios with and without inter-cellular interference, and without the need for additional overhead. We also demonstrated how the proposed method outperforms other similar protocols, particularly in extremely congested scenarios,i.e., $K_0 > 25000$. The robustness of the proposed method has also been evaluated by varying the number of BS antennas, demonstrating that the results are still valid for a number of antennas other than those considered during the training phase of the proposed NN-based method. The robustness of the proposed method has also been assessed regarding the number of BS antennas and edge SNR, demonstrating that the proposed method is a promising GB RA protocol option.

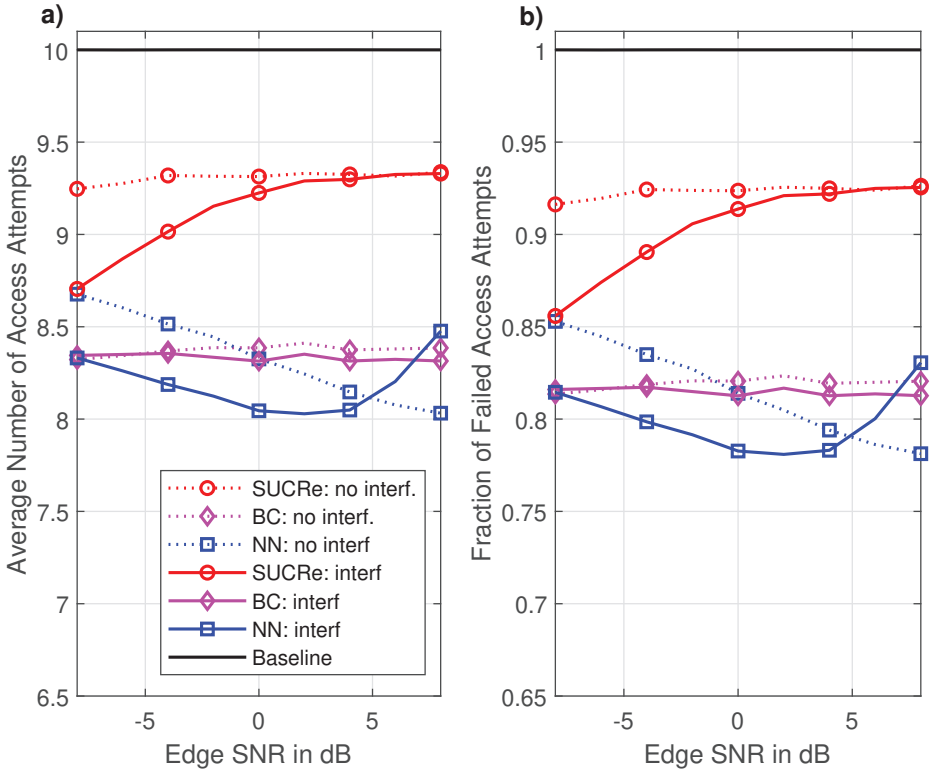


Figure 22 – Performance with edge SNR variation in dB: a) ANAA, and b) FFAA. With $M = 100$, $\tau_p = 10$, $K_{\text{ici}} = 10$ and $K_0 = 28000$.

5 APPLICATION OF Q-LEARNING ALGORITHM FOR PILOT COLLISION RESOLUTION IN GRANT-FREE RANDOM ACESS SYSTEMS

This chapter presents a new GF RA protocol which we denote by massive QmMTC (mQmMTC). This protocol is a collaborative multi-agent QL-based method applied to pilot collision resolution in mMTC systems, similar to the protocol proposed in (SHARMA; WANG, 2019). However, unlike the introduced in (SHARMA; WANG, 2019), the mQmMTC is designed to work in massive MIMO systems under realistic propagation effects, such as multipath fading, shadowing, path loss, thermal noise and ICI.

5.1 EXISTING QL-BASED GF RA PROTOCOLS

Q-learning can be used as a multi-agent RLe method to assist IoT devices in selecting the least congested RA pilots in a decentralized manner. In typical mMTC networks, active IoT devices can coordinate neither with the BS nor with each other for pilot selection. Thus, IoT devices act as individual learning agents, using their previous experience to increase the probability of selecting exclusive RA pilots and reducing the occurrence of pilot collisions. In this section, we revisit two earlier works in which this framework was applied in an *simple collision channel*, *i.e.*, assuming that the only communication impairment is *pilot collision*.

We consider a mMTC network employing a GF RA policy. The cell consists of K_a IoT active devices, each with L packets to transmit competing for τ_p pilot resources. After transmitting all L packets, the k-th device becomes inactive, and a new device with more L packets to transmit activates and takes the position of the inactive one, ensuring that the cell always has K_a active devices.

The devices send the payload data prepended with a RA pilot to enable channel estimation using a GF RA protocol. Suppose the RA pilot is chosen exclusively by that device (no interference). In that case, we assume that the payload data packet, transmitted by the device, is successfully decoded at the BS, and the device proceeds to transmit the next data packet of a total of L packets. Otherwise, if a collision occurs, the device repeats the transmission of the same packet in the next frame, increasing *latency* and decreasing network *throughput*. Thus, RLe techniques can be employed to guide the pilots' choice of devices towards the least congested ones, improving the connectivity performance of the network as a whole. In this work, the performance metric *latency* is normalized and expressed as multiples of time slot, T_s. We define

latency as the total amount of time spent on transmitting L packets, including failed transmission attempts, divided by the average duration of a time slot, T_s. If transmitting 1 packet takes T_s seconds and each failed attempt also takes T_s seconds, then the latency is expressed as:

$$\textit{latency} = \frac{L \times T_s + A_F \times T_s}{T_s} = (L + A_F) \quad [T_\mathrm{s}] \tag{73}$$

where A_F is the number of failed packet transmission attempts. The normalization procedure ensures that the latency results are independent of the duration of T_c, which varies depending on the environment and device's velocity as discussed in subsection (2.3.4).

The interaction between the IoT device and the environment can be modeled as a Markov decision process (MDP), where at each time step, a device can change its current state $s_t \in S$ to $s_{t+1} \in S$ by taking action $a_t \in A$ based on a transition probability function $f(s_t,a_t,s_{t+1})$ (SHARMA; WANG, 2019). Depending on its state-action pair, the device is rewarded with $r_{t+1} \in \mathcal{R}$ during the transition. Besides, the expected return of a state-action pair is given by $Q^\pi(s,a) = \mathbb{E}\left[\sum_{j=0}^{E_p} \gamma^j r_{t+j+1} | s_t = s, a_t = a, \pi\right]$, where π is the established policy, $\gamma \in [0,1]$ is the discount factor, and E_p is the length of one episode (SHARMA; WANG, 2019).

Satisfying the Bellman optimality equation, the Q-function can then be written as $Q^*(s,a) = \max_\pi Q^\pi(s,a)$. If a *greedy policy* $\pi(s) = \arg\max_a Q^\pi(s,a)$ is established for the Q-function, then we have a Q-Learning algorithm that selects only the actions with the highest Q-value $Q(s,a)$ at each state, calculated iteratively as:

$$Q_{t+1}(s_t,a_t) = Q_t(s_t,a_t) \; + \kappa_t \left[r_{t+1} + \gamma \max_a Q_t(s_{t+1},a_{t+1}) - Q_t(s_t,a_t) \right], \tag{74}$$

where κ_t is the learning rate at the tth time step. This model can then be applied for *decentralized pilot selection* as described in (BELLO *et al.*, 2014),(SHARMA; WANG, 2019):

$$Q_{t+1}(k,\ell) = Q_t(k,\ell) + \kappa \left[R_F(k,\ell) - Q_t(k,\ell) \right], \tag{75}$$

where $Q(k,\ell)$ indicates that the k-th device has chosen the pilot $\ell = c(k)$ and $R_F(k,\ell)$ indicates the reward function for this choice.

In (BELLO *et al.*, 2014), a simple method for computing the rewards $R_F(k,\ell)$ in (75) is proposed via an *independent* Q-learning approach for mMTC (**iQmMTC**), where $R_F(k,\ell) = +1$ if the transmission succeeds, or $R_F(k,\ell) = -1$ otherwise. However, (SHARMA; WANG, 2019) has demonstrated that better results can be obtained through a *collaborative* approach (**cQmMTC**), in which the device is rewarded either with $R_F(k,\ell) = +1$ for a successful

transmission or with $R_F(k,\ell) = -P_c(k)$ if the transmission fails, with the collaborative penalty function $P_c(k)$ being computed as:

$$P_c(k) = \frac{1}{K_a} C_L(k), \tag{76}$$

where K_a is the number of active devices in the cell and the congestion level $C_L(k) = |\mathcal{S}_{c(k)}|$ is the *number of contenders* for $c(k)$ (SHARMA; WANG, 2019).

We assume that each device maintains its own $1 \times \tau_p$ Q-table. Initially, all K_a devices have zeroed Q-tables, and all τ_p RA pilots are equally likely to be chosen by any of them. Then, on each subsequent transmission attempt, all devices are rewarded with $R_F(k,\ell)$, and their Q-tables are updated using (75). Once a device's Q-table has been updated, it will only be able to select pilots with the highest Q-value. The process is repeated until the L packets are sent. It is worth noting that the new devices activated with probability P_a also begin transmission attempts with a zeroed Q-table. As a result, it must also update its Q-table after each transmission attempt in order to learn from the current set environment.

In the cQmMTC approach of (SHARMA; WANG, 2019), the congestion levels in the event of a collision are computed as the number of contending devices, $|\mathcal{S}_{c(k)}|$. However, nothing is said about the feasibility of making this information available on the device's side. At first glance, it appears that it would necessitate developing an estimator to be used at the BS and then feeding the results back to the devices, which would incur significant signaling overhead. To avoid this and simplify the procedure, we propose a 2-step Q-learning-based GF RA protocol which also takes advantage of a large number of BS antennas to allow a collaborative penalty function computation at the device's side in case of collisions with minimal complexity and overhead.

5.2 PROPOSED 2-STEP Q-L GF RA M-MIMO PROTOCOL

This section describes the proposed 2-step, GF, RA *Q-Learning* based pilot collision resolution protocol.

Let $\mathcal{K}$ be the set of devices in the cell, which are activated with probability P_a. These devices transmit not only *payload data* but also a randomly selected *pilot sequence* during the UL to enable channel estimation at the BS side. We consider τ_p mutually orthogonal pilot sequences: $\{\boldsymbol{\psi}_1,...,\boldsymbol{\psi}_{\tau_p}\} \in \mathbb{C}^{\tau_p \times 1}$, such that each pilot has length τ_p and squared magnitude $\|\boldsymbol{\psi}_t\|^2 = \tau_p$, $\forall t \in [1, \tau_p]$. We also consider a TDD scheme, in which the channels are considered constant

during a time slot. The BS is equipped with M antennas localized at the center of the cell. An illustrative representation of the adopted scenario is presented in Figure 23.

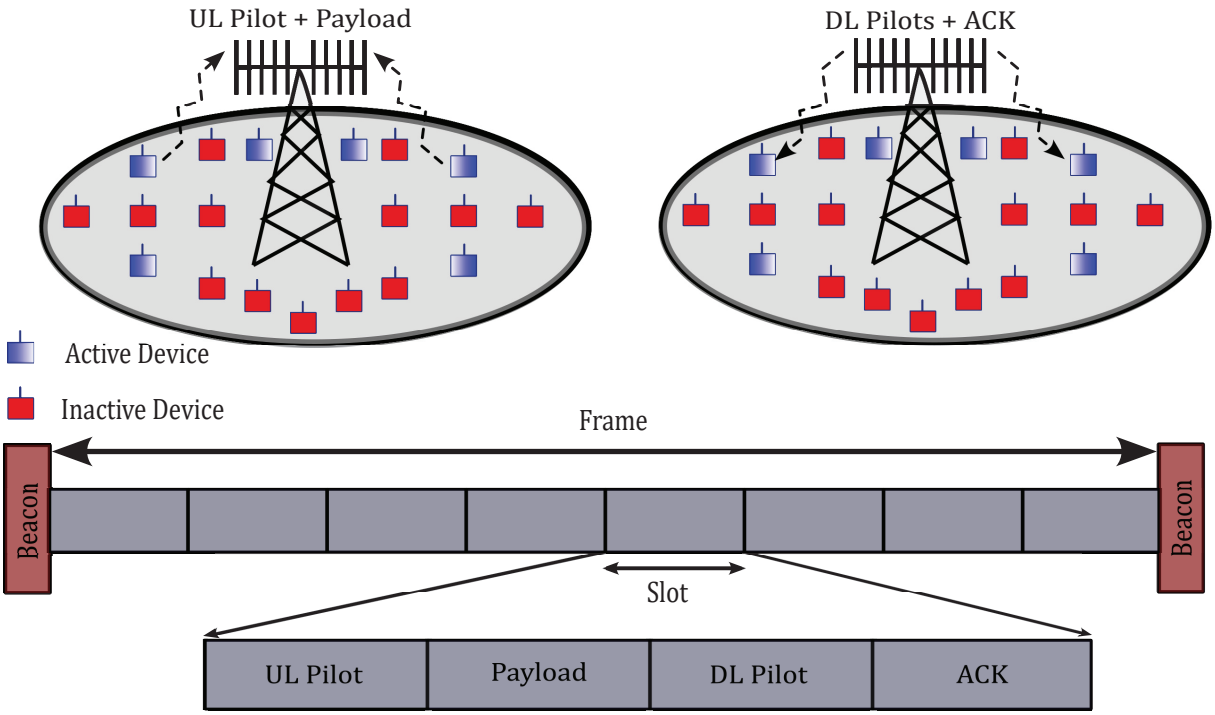


Figure 23 – Illustrative representation of the adopted scenario.

Each device that activates has a number of L_k packets to transmit. Therefore, considering $\mathcal{S}_t \subset \mathcal{K}$ as the set of devices that want to transmit data selecting pilot t, its cardinality follows a binomial distribution:

$$|\mathcal{S}_t| \sim \mathcal{B}\left(K_T, \frac{P_a}{\tau_p}\right) \tag{77}$$

where $K_T = |\mathcal{K}|$ is the total number of devices in the cell and P_a is their activation probability. The channel vector between BS and device k is denoted by $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ and follows a complex Gaussian distribution $\mathbf{h}_k \sim \mathcal{CN}(0, \beta_k \mathbf{I}_M)$, where β_k is the large-scale fading coefficient, which follows an urban micro scenario (3GPP, 2018). The large-scale fading of the link between device k and its BS is calculated as:

$$\beta_k = 10^{-\vartheta \log(dist_k) + \frac{P_L + \varphi}{10}}. \tag{78}$$

where $dist_k$ is the distance between device k and the BS, $\vartheta = 3.8$ is the path loss exponent, $\varphi \sim \mathcal{N}(0, \sigma_{sf}^2)$ is the shadow fading, with standard deviation $\sigma_{sf} = 10$ dB, and $P_L = -34.53$ dB is the path loss at the reference distance (BJÖRNSON *et al.*, 2017a; NISHIMURA *et al.*, 2020).

When the device k transmits data, it randomly selects one of the τ_p pilot sequences and transmits it followed by its UL payload data $\mathbf{d}_k \in \mathbb{C}^{\tau_d \times 1}$, with a non-zero transmit power

$\rho_k > 0$. We can denote the chosen pilot as $c(k) \in \{1,2,\ldots,\tau_p\}$, and define the UL signal as $\mathbf{x}_k = [\boldsymbol{\psi}_{c(k)}^T, \mathbf{d}_k^T]^T \in \mathbb{C}^{(\tau_p+\tau_d)\times 1}$.

Thus, the BS receives the signal

$$\mathbf{Y} = [\mathbf{Y}_p, \mathbf{Y}_d] = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k\mathbf{x}_k^T + \mathbf{N}, \tag{79}$$

where $\mathbf{Y} \in \mathbb{C}^{M\times(\tau_p+\tau_d)}$, $\mathbf{Y}_p \in \mathbb{C}^{M\times\tau_p}$, $\mathbf{Y}_d \in \mathbb{C}^{M\times\tau_d}$, and $\mathbf{N} \in \mathbb{C}^{M\times(\tau_p+\tau_d)}$ is the receiver noise with entries drawn from $\mathcal{CN}(0,\sigma^2)$. Besides, we have that

$$\mathbf{Y}_p = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k\boldsymbol{\psi}_{c(k)}^T + \mathbf{N}_p, \ \textit{and} \tag{80}$$

$$\mathbf{Y}_d = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k\mathbf{d}_k^T + \mathbf{N}_d, \tag{81}$$

with $\mathbf{N} = [\mathbf{N}_p, \mathbf{N}_d]$. Hence, the BS correlates (80) with each pilot to generate channel estimates. For the case of an arbitrary pilot $\boldsymbol{\psi}_t$, with $t \in [1, \tau_p]$, it yields:

$$\mathbf{y}_t = \mathbf{Y}_p\frac{\boldsymbol{\psi}_t^*}{\|\boldsymbol{\psi}_t\|} = \sum_{i\in\mathcal{S}_t} \sqrt{\rho_i\tau_p}\mathbf{h}_i + \mathbf{n}_t, \tag{82}$$

where $\mathbf{n}_t = \mathbf{N}_p\frac{\boldsymbol{\psi}_t^*}{\|\boldsymbol{\psi}_t\|}$ is the effective receiver noise, so that $\mathbf{n}_t \sim \mathcal{CN}(0,\sigma^2\mathbf{I}_M)$. As a result, the BS tries to decode the payloads in (81) using the channel estimates $\mathbf{y}_t$, evaluating:

$$\widehat{\mathbf{d}}_k^T = \frac{\mathbf{y}_t^H}{\sqrt{\tau_p}}\mathbf{Y}_d. \tag{83}$$

The signal-to-interference-plus-noise ratio (SINR) of $\widehat{\mathbf{d}}_k$ in (83) can be obtained following the SINR analysis of (MARINELLO *et al.*, 2022), adapting the results to our scenario.

First, the components of $\widehat{\mathbf{d}}_k^T$ are expanded as:

$$\begin{aligned}\widehat{\mathbf{d}}_k^T &= \frac{1}{\sqrt{\tau_p}}\left[\sum_{i\in\mathcal{S}_t}\sqrt{\rho_i\tau_p}\mathbf{h}_i + \mathbf{n}_t\right]^{\mathrm{H}}\left[\sum_{k\in\mathcal{K}}\sqrt{\rho_k}\mathbf{h}_k\mathbf{d}_k^T + \mathbf{N}_d\right], \\ &= \underbrace{\rho_k\|\mathbf{h}_k\|^2\mathbf{d}_k^T}_{\omega_1} + \underbrace{\sum_{i\in\mathcal{S}_t, i\neq k}\rho_i\|\mathbf{h}_i\|^2\mathbf{d}_i^T}_{\omega_2} + \underbrace{\sum_{i\in\mathcal{S}_t}\sum_{k\in\mathcal{K},k\neq i}\sqrt{\rho_i\rho_k}(\mathbf{h}_i)^{\mathrm{H}}(\mathbf{h}_k)\mathbf{d}_k^T}_{\omega_3} + \\ &\quad \underbrace{\frac{(\mathbf{n}_t)^{\mathrm{H}}}{\sqrt{\tau_p}}\sum_{k\in\mathcal{S}_t}\sqrt{\rho_k}\mathbf{h}_k\mathbf{d}_k^T}_{\omega_4} + \underbrace{\frac{(\mathbf{y}_t)^{\mathrm{H}}}{\sqrt{\tau_p}}\sum_{k\in\mathcal{K}\setminus\mathcal{S}_t}\sqrt{\rho_k}\mathbf{h}_k\mathbf{d}_k^T}_{\omega_5} + \underbrace{\frac{(\mathbf{y}_t)^{\mathrm{H}}}{\sqrt{\tau_p}}[\mathbf{N}_d]}_{\omega_6}.\end{aligned} \tag{84}$$

Next, the expected power of each term in (84) is computed as the expectation of their squared magnitude, following the analysis in (MARINELLO *et al.*, 2017). Table 15 shows the

mean and variance of each component in (84). Then, taking only $\mathbb{E}[\omega_1]$ as the desired signal power, while relegating $\mathbb{V}[\omega_1]$ as an additional interference as in (MARINELLO *et al.*, 2017), the full SINR expression is obtained as γ_k^{ul} in (85):

Table 15 – First and Second-Order Statistical Moments of the decoded signal

Term ω_i	$\mathbb{E}[\omega_i]$	$\mathbb{V}[\lvert\omega_i\rvert]$
ω_1	$M\rho_k\beta_k$	$M\rho_k\beta_k$
ω_2	$M\sqrt{\sum_{i\in\mathcal{S}_t, i\neq k}\rho_i^2\beta_i^2}$	$M\sum_{i\in\mathcal{S}_t, i\neq k}\rho_i\beta_i$
ω_3	0	$M\sum_{i\in\mathcal{S}_t}\sum_{k\in\parallel}\rho_i\rho_k\beta_i\beta_k$
ω_4	0	$\frac{M\sigma^2}{\tau_p}\sum_{k\in\mathcal{S}_t}\rho_k\beta_k$
ω_5	0	$M[\sum_{k\in\mathcal{S}_t}\rho_i\beta_i+\frac{\sigma^2}{\tau_p}]\sum_{k\in\mathcal{K}\setminus\mathcal{S}_t}\rho_k\beta_k$
ω_6	0	$M[\sum_{k\in\mathcal{S}_t}\rho_i\beta_i+\frac{\sigma^2}{\tau_p}]\sigma^2$

$$\gamma_k^{\text{ul}} = \frac{M\,\rho_k^2\,\beta_k^2}{M\sum_{\substack{i\in\mathcal{S}_t\\ i\neq k}}\rho_i^2\,\beta_i^2 + \left[\sum_{i\in\mathcal{S}_t}\rho_i\beta_i + \frac{\sigma^2}{\tau_p}\right]\left[\sum_{j\in\mathcal{K}}\rho_j\beta_j + \sigma^2\right]}. \tag{85}$$

We assume that the decoding of (83) is always successful when k is the unique competitor for the pilot t (No pilot collisions occur). The BS responds with an ACK feedback message if the decoding of (83) is successful, together with the transmission of a precoded DL pilot signal $\mathbf{V}\in\mathbb{C}^{M\times\tau_p}$, with power q, according to:

$$\mathbf{V} = \sqrt{\frac{q}{\tau_p}}\sum_{t=1}^{\tau_p}\frac{\mathbf{y}_t^*}{\|\mathbf{y}_t\|}\boldsymbol{\psi}_t^T. \tag{86}$$

The devices receive $\mathbf{z}_k\in\mathbb{C}^{\tau_p\times 1}$, $k\in\mathcal{S}_t$

$$\mathbf{z}_k^T = \mathbf{h}_k^T\mathbf{V} + \boldsymbol{\eta}_k^T, \tag{87}$$

where $\boldsymbol{\eta}_k\sim\mathcal{CN}(\mathbf{0},\sigma^2\mathbf{I}_{\tau_p})$ is the noise. After correlating $\mathbf{z}_k$ with $\boldsymbol{\psi}_t$, the device calculates

$$z_k = \mathbf{z}_k^T\frac{\boldsymbol{\psi}_t^*}{\|\boldsymbol{\psi}_t\|} = \sqrt{q}\,\mathbf{h}_k^T\frac{\mathbf{y}_t^*}{\|\mathbf{y}_t\|} + \eta_k, \tag{88}$$

where $\eta_k\sim\mathcal{CN}(0,\sigma^2)$.

Collaborative penalty function computation. Let $\alpha_t = \sum_{i\in\mathcal{S}_t}\rho_i\beta_i\tau_p$ be the *sum of channel gains of the devices* in $\mathcal{S}_t$ seen at the BS according to (82), then an asymptotically *error-free estimator* for α_t is proposed in a similar scenario[1] in (BJÖRNSON *et al.*, 2017a) as:

$$\hat{\alpha}_{t,k} = \max\left(\left[\frac{\Gamma(M+\frac{1}{2})}{\Gamma(M)}\right]^2\frac{q\rho_k\beta_k^2\tau_p}{[\Re(z_k)]^2} - \sigma^2,\ \rho_k\beta_k\tau_p\right), \tag{89}$$

[1] In Björnson *et al.* (2017a), the $\hat{\alpha}_{t,k}$ estimate is computed as part of the 4-steps GB handshake procedure of SUCRe protocol, and used to let the UEs decide whether they should retransmit the chosen pilot or not, depending if it is the strongest contender. Differently, we employ here a GF RA protocol, in which the devices transmit the payload data together with an RA pilot to enable channel estimation in a reduced number of steps.

where $\Re(\cdot)$ is the real part and $\Gamma(\cdot)$ is the complete Gamma function.

Based on the estimate $\hat{\alpha}_{t,k}$ in (89), reminded that $\alpha_{c(k)} = \sum_{i \in \mathcal{S}_{c(k)}} \rho_i \beta_i \tau_p$, and since the device k is aware of its average channel gain β_k, the device can calculate a measure of the congestion level of its selected pilot as follows:

$$\widehat{\phi}_k = \frac{\hat{\alpha}_{t,k}}{\rho_k \beta_k \tau_p}, \qquad 1 \leq \phi_k < \infty. \tag{90}$$

The large-scale fading coefficient β_k can be estimated from the broadcast signals periodically received from the BS, such as beacon[2] signals. One can note that as $\widehat{\phi}_k$ approaches 1, it indicates that it is likely that no other device has chosen the pilot $c(k)$. On the other hand, as $\widehat{\phi}_k$ increases, it indicates that it is likely that many other devices chose the same pilot $c(k)$. Therefore, $\widehat{\phi}_k$ can be seen as a *rough estimate* of $C_L(k)$; hence, the penalty function $P_c(k)$ in (76) can be computed approximately as

$$P_c(k) \approx \frac{\widehat{\phi}_k}{\widehat{K}_a}, \tag{91}$$

in which $\widehat{K}_a$ is an estimate for the number of active devices in the cell. We propose to use a simple estimator that calculates $\widehat{K}_a$, the expected number of active devices, assuming no pilot collision occurs and that the device k knows the total number of users K_T in the center cell, the activation probability P_a, and the average number of packets $\mathbb{E}[L_k]$ transmitted to the BS by each active device in the cell. In this way, one can compute:

$$\widehat{K}_a = K_T \cdot P_a \cdot \mathbb{E}[L_k]. \tag{92}$$

The penalty function $P_c(k)$ in (91) can be used in crowded networks, considering realistic propagation effects such as multipath fading, path loss, thermal noise, and ICI, allowing practical implementation of the cQmMTC approach in a GF RA protocol. Therefore, in the proposed protocol, the devices choose their RA pilots along the transmissions of the L_k packets according to their own Q-table, which is updated following (75), while computing the rewards as:

$$R_F(k,\ell) = \begin{cases} +1, & \text{if the transmission succeeds} \\ -P_c(k) = -\widehat{\phi}_k / \widehat{K}_a, & \text{in case of pilot collision.} \end{cases} \tag{93}$$

The diagram presented in Figure 24 illustrates K devices with L packets to transmit while competing for τ_p pilots. It also shows the rewards estimated by each device with or without

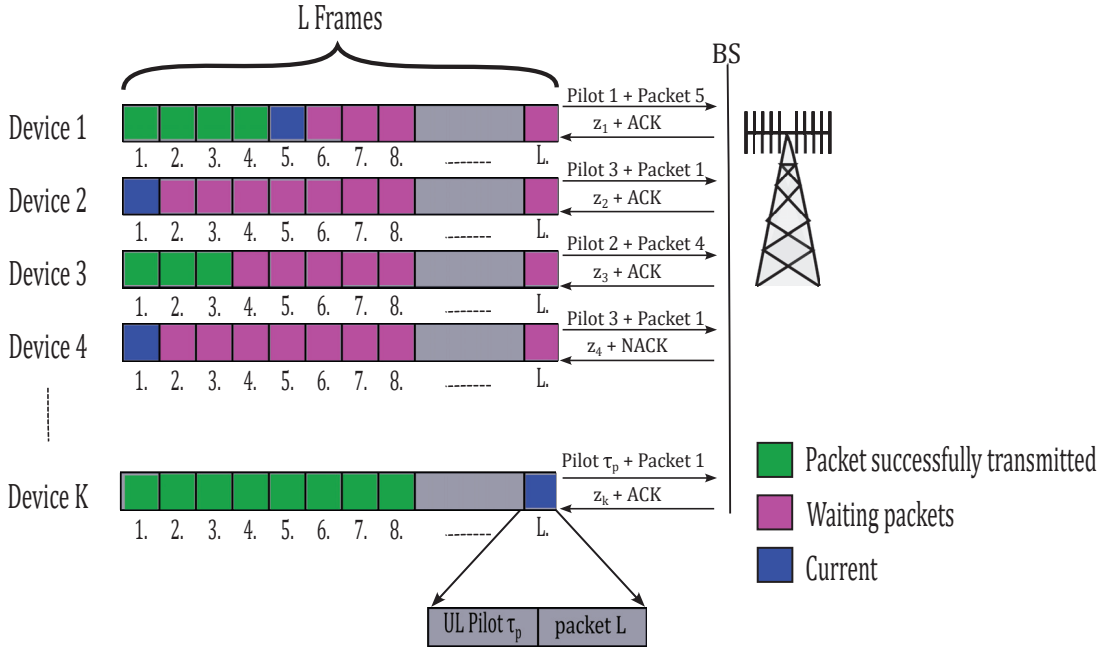
[2] The beacon signal is a periodically transmitted signal by the BS and carries important information such as the cellular network identifier, timestamp, gateway address, paging area ID, and other BS parameters. This information can be used by devices to estimate their own channel gain (SJ *et al.*, 2001).

collisions. Figure 24.a illustrates that each device has L packets to transmit. The green squares represent the packets that are already successfully transmitted, the magenta squares represent the packets waiting to be sent in the next RA block, and the blue squares represent the packets currently being transmitted. Each packet is contained within a frame with the UL pilot chosen by the device. Based on the pilot chosen by the device, the BS responds with an ACK feedback if the transmission succeeds or with a NACK if the transmission fails. In this work, the NACK response is interpreted simply as the absence of an ACK response. In all cases, successful or unsuccessful, the devices receive z_k, which allows them to estimate their congestion levels. In Figure 24.b, the columns represent the available pilots, and the rows represent each active device. The pilot chosen by a particular device is depicted in colored rectangles as green or red. Green rectangles indicate that the transmission was successful (no pilot collision occurred), and red rectangles indicate that the packet transmission was unsuccessful (a pilot collision occurred). The current packet being transmitted is also expressed in all colored rectangles. Finally, Figure 24.c represents how each device calculates its reward based on the occurrence or not of pilot collision and associates this reward with its chosen pilot in the current RA block.

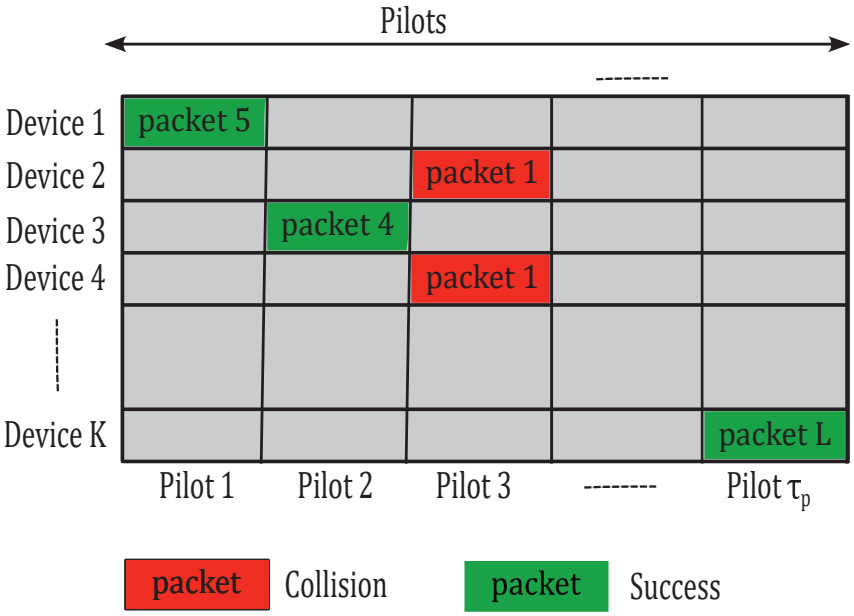
5.3 NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed QL-based GF RA protocol in terms of (**i**) average latency, (**ii**) average network throughput, defined as the ratio between the number of successfully transmitted packets (without collisions) at certain time step and the number of available pilots τ_p, and (**iii**) per-user throughput, considered as the ratio of the total number L_k of successfully sent packets by each device to the total number of transmission attempts, A_T, that the device has to make to send them, in such a way that $L_k \leq A_T$. For the simulations, we consider a massive MIMO BS equipped with M antennas at the center of a hexagonal cell with a radius of R, surrounded by six neighboring hexagonal cells with the same radius. The simulation parameters are shown in Table 16. The description column of the Table indicates Figures that depict the results of simulations that employ the respective parameter value. Parameters without indication of Figures are applied in all simulations of the section.

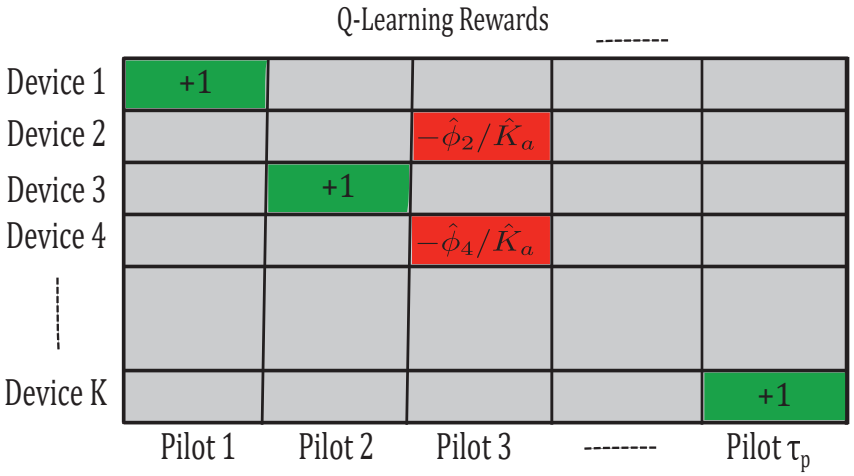
With respect to L and P_a, we investigate three different scenarios in this section: (**i**) $L_k = L, \forall k \in \mathcal{K}$, and $P_a = 1$, such that $K_a = K_T$, that is, the number of active devices in the center cell is equal to its total number of devices since all devices are active; (**ii**) random L_k, and $P_a = 1$, also resulting in $K_a = K_T$; (**iii**) random L_k, $\forall k \in \mathcal{K}$, and $P_a = 0.1\%$, such that



(a) K devices with L packets to transmit;



(b) K devices competing for τ_p pilots;



(c) Q-Learning rewards

Figure 24 – Illustrative representation of the proposed protocol.

Table 16 – Simulation parameters of the proposed QL based GF RA protocol

Parameter	Value	Description
M	100	Number of BS antennas in the center and neighboring cells. Figures (25, 26, 30,32,33,34, 35,36)
M	from 1 to 100	Variation in the number of BS antennas in the center and neighboring cells. Figures (27 and 28)
P_a	1	Probability of the devices activation. Figures (25, 26, 33 and 34)
P_a	0.001	Probability of the devices activation. Figures (35 and 36)
τ_p	400	Number of available RA pilot sequences. Figures (25, 26,27,28, 29, 30,32,33,34)
τ_p	40	Number of available RA pilot sequences. Figures (35, 36)
ρ	27 dBm	Transmit power of UEs
q	27 dBm	Transmit power of the BS
σ^2	-98.65 dBm	Noise variance
$K_{\mathrm{i}ci}$	K_a, $10 \cdot K_a$, $60 \cdot K_a$	Number of active devices in the neighboring cells. Figures (25, 26, 30, 31, 32, 33, 34)
$K_{\mathrm{i}ci}$	400	Number of devices in the neighboring cells. Figures (27, 28,29)
$K_{\mathrm{i}ci}$	40	Number of available RA pilot sequences. Figures (35, 36)
K_T	from 0 to 800	Number of devices in center cell. Figures (25, 26, 33, 34)
K_T	400	Number of devices in the center cell. Figures (27, 28, 29)
K_T	600	Number of devices in the center cell. Figures (30, 32)
K_T	from 0 to 4000	Number of devices in the center cell. Figures (35, 36)
L_k	10	Number of packets that user k have to transmit. Figures (25, 26, 27,28, and 29)
L_k	from 1 to 100	(Figures 30,31 and 32)
L_k	Random with distribution $\mathcal{U}(1,10)$	Number of packets that user k has to transmit. Figures (33, 34, 35, 36)
R	250 m	Radius of the center cell
R	250 m	Radius of the neighboring cells
$\sigma_{\mathrm{s}f}$	10 dB	Shadow fading standard deviation
P_L	- 34.53 dB	Path loss at reference distance.
ϑ	3.8	Path loss exponent.
	5G sub-6 GHz	Band of operation
	27 dBm	Transmit power of devices in adjacent cells

$K_a \leq K_T$ is also random.

We investigate 4 protocols: the **a**) *baseline* scheme, which is equivalent to the slotted ALOHA protocol, with the devices choosing the pilots uniformly at random; **b**) *iQmMTC* approach of (BELLO *et al.*, 2014); **c**) *cQmMTC* approach of (SHARMA; WANG, 2019), assuming that the actual values of $|\mathcal{S}_t|$ and K_a are perfectly known at the devices' side, like if a genie could inform this to them; and **d**) our proposed 2-step QL-based GF RA protocol, which takes advantage of the *massive* MIMO propagation features to efficiently compute the negative rewards of the QL framework at the devices' side, which we denote as *mQmMTC*. Besides, for this last one, we investigate its performance in the scenarios with and without ICI[3]. The results *mQmMTC* with ICI in Figures 25, 26, 30, 31, 32, 33 and 34 are presented with three levels of

[3] It is worth noting that for the baseline, iQmMTC, and cQmMTC in the adopted scenario, only pilot collisions degrade their connectivity performance. Therefore, ICI does not matter for them.

K_{ici} values, $K_{\text{ici}} = K_a$, $K_{\text{ici}} = 10 \cdot K_a$, and $K_{\text{ici}} = 60 \cdot K_a$. Figures 27, 28, 29 present results with a fixed number of $K_{\text{ici}} = 400$ and Figures 35, 36 depict results with $K_{\text{ici}} = 40$.

5.3.1 Fixed L_k and $P_a = 1$

In this subsection, we take $L_k = L, \forall k \in \mathcal{K}$, and $P_a = 1$, such that $K_a = K_T$. Although these simplifying assumptions usually do not hold in practice, they are useful to unveil the full potential of the investigated methods and evaluate the performance losses when not assuming them, which is carried out in the following subsections. The graphs presented in Figures 25 and 26 have been generated with 10000 Monte-Carlo realizations. Each realization is a frame, or a time step in the QL framework, in which each device can transmit only one pilot and the payload packet with a number of devices varying from 25 to 800.

Figure 25 displays the average latency, in multiples of time-slots T_s, as a function of K_a results. The closeness of the proposed mQmMTC results to the ideal cQmMTC protocol, in both scenarios with and without interference and for the three levels of K_{ici}, is remarkable. Additionally, both the mQmMTC and ideal cQmMTC results are lower than the baseline for all K_a values. The results are below the independent QmMTC for $K_a > 400$, confirming the superiority of cQmMTC as stated in (SHARMA; WANG, 2019). It is noteworthy that even scaling the number of devices in the neighboring cell to $K_{\text{ici}} = 60 \cdot K_a$ is not enough to impair the functionality of the proposed method.

Similarly, Figure 26 reveals the average network throughput versus K_a results, where the performance of proposed mQmMTC protocol is close to the ideal cQmMTC performance for both the scenarios without and with ICI, considering $K_{\text{ici}} = K_a$, while being consistently superior to the baseline results for any number of devices K_a and any investigated values of K_{ici}. Furthermore, compared with iQmMTC, the obtained performances are very similar in the region of $K_a \leq \tau_p = 400$ devices, while the performance obtained by the proposed mQmMTC approach becomes remarkably superior for a higher number of devices. For example, with $K_a = 600$ active devices, the baseline and iQmMTC results achieve average network throughputs of ≈ 0.33 and 0.35, respectively, and our proposed mQmMTC protocol achieves an average network throughput of ≈ 0.43 in the worst case scenario with ICI and $K_{\text{ici}} = 60 \cdot K_a$, an improvement of $\approx 30\%$ and $\approx 22\%$, respectively.

The graphs presented in Figure 27 and 28 are generated with 80000 Monte-Carlo realizations. The number of devices in the center and neighboring cells are fixed at $K_a = K_{\text{ici}} =$

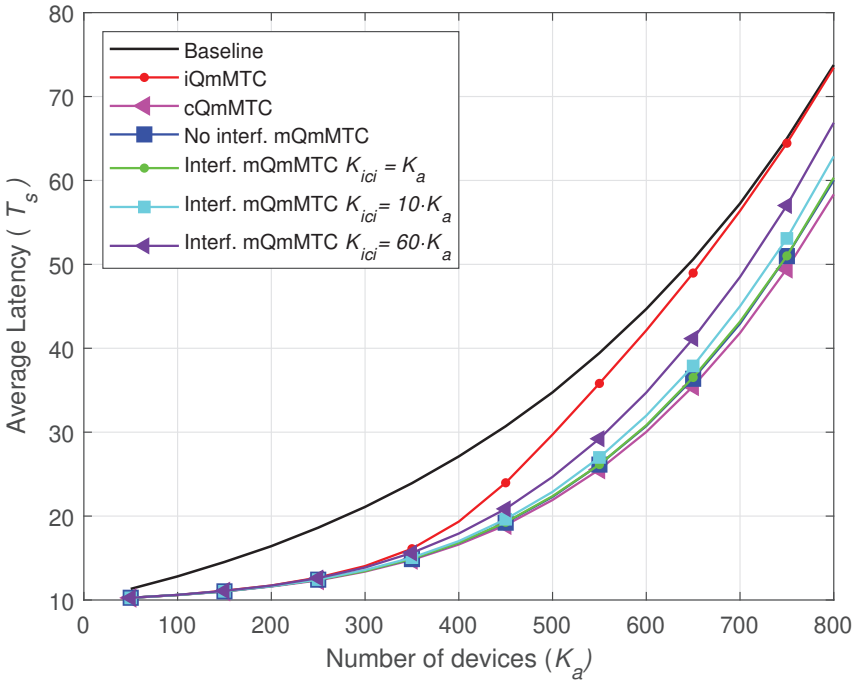


Figure 25 – Average latency $\times K_a$, for $L = 10$ packets, $P_a = 1$, $\tau_p = 400$, $K_{\text{ici}} = K_a$, $10 \cdot K_a$, $60 \cdot K_a$, and $M = 100$.

400 and the number of available pilots is also fixed at $\tau_p = 400$. The number of BS antennas varies in the range $M \in [1, 100]$, both in the center and neighboring cells. The increase of antennas in the neighboring cells is carried out only for computational reasons. In practice, it does not impact the results and does not differ from the results of the increase only in the central cell since only the UL time is considered in this analysis.

Figure 27 depicts the average latency with an increasing number of BS antennas M and Figure 28 reveals the behavior of the average network throughput *vs.* M. Both Figures of merit for the proposed mQmMTC approach improve with the increasing number of antennas M, since the reward computations in (91) benefit from the large number of BS antennas (favorable propagation effect). The results of the ideal cQmMTC are also included in the Figures as a lower bound (avg. latency) and upper bound (avg. throughput), respectively. The average percentual degradations of the results of mQmMTC regarding the ideal cQmMTC is also shown in both Figures for $M = 30$ and $M = 100$. One can see that the improvement caused by increasing M from $M = 10$ to $M = 100$ is not as significant as when increasing M from $M = 1$ to $M = 10$. Therefore we can conclude that our proposed mQmMTC RA protocol is able to achieve very improved connectivity performance even with a small number of BS antennas.

Figure 29 shows the throughput and latency degradation levels versus the number of antennas M. Indeed, the graphs reaveals that $M \approx 30$ antennas at the BS are sufficient to attain

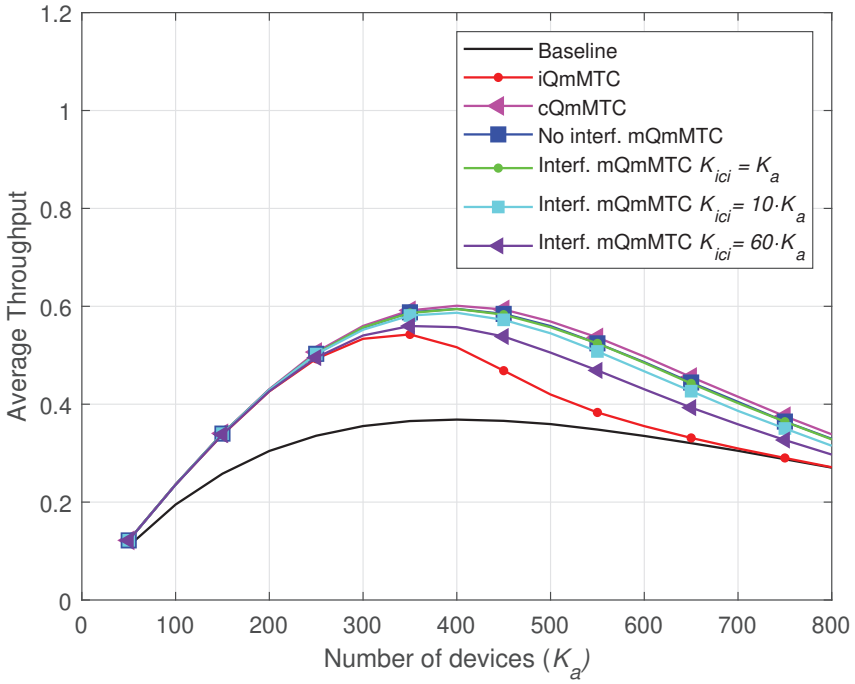


Figure 26 – Average network throughput $\times K$**, for** $L = 10$ **packets,** $P_a = 1$**,** $\tau_p = 400$**,** $K_{\text{ici}} = K_a$**,** $10 \cdot K_a$**,** $60 \cdot K_a$**, and** $M = 100$**.**

a reliable congestion level estimation of $\widehat{\phi}_k$, when a maximum acceptable level of throughput degradation of 3% is considered.

Figures 30, 31 and 32 present, respectively, the graphs of latency, normalized latency, which is the latency normalized by L, and network throughput *versus* L . The results shown in

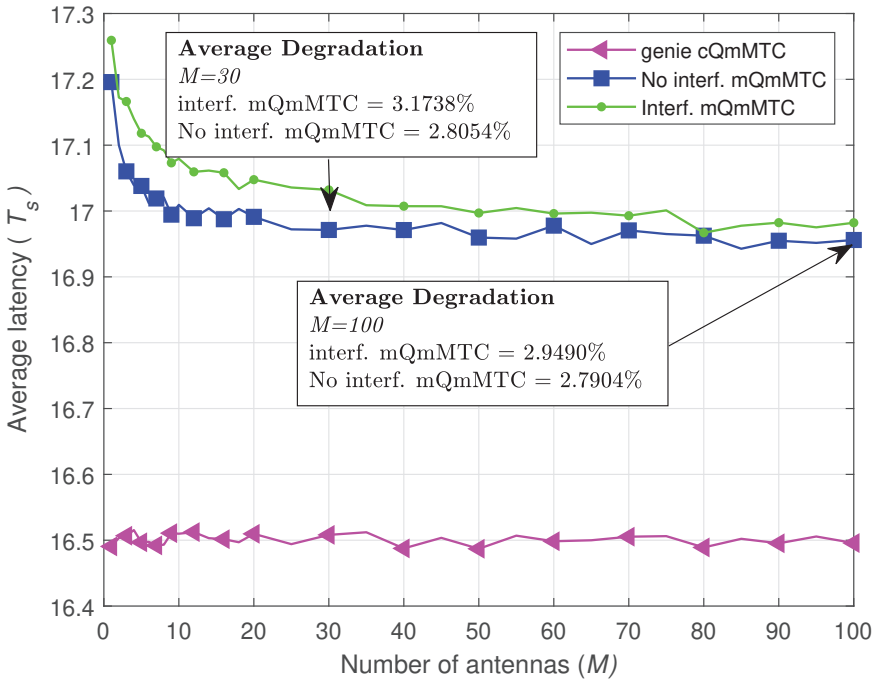


Figure 27 – Average latency $\times M$**, for** $L = 10$ **packets,** $\tau_p = 400$**,** $K_{\text{ici}} = 400$**, and** $K_a = 400$**.**

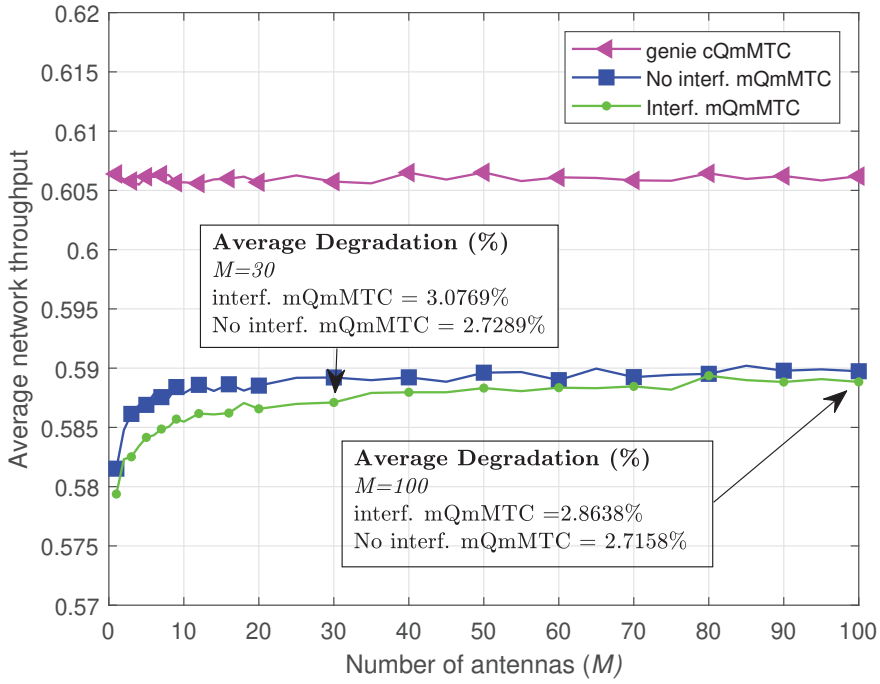


Figure 28 – Average network throughput $\times M$, for $L = 10$ packets, $\tau_p = 400$, $K_{\text{ici}} = 400$, and $K_a = 400$.

both Figures are generated with 10000 Monte-Carlo realizations. The number of active devices are kept fixed at $K_a = 600$, the number of antennas is also fixed at $M = 100$ and the three K_{ici} levels, $K_{\text{ici}} = K_a$, $K_{\text{ici}} = 10 \cdot K_a$ and $K_{\text{ici}} = 60 \cdot K_a$ are also investigated. The graphs illustrate the superiority of the proposed methods even when each device has a small number of packets to send ($L \leq 10$). In fact, a minimum number of $L = 2$ is enough for the proposed methods (mQmMTC without and with ICI) to produce a result superior to the baseline and the iQmMTC methods, while approximating to the cQmMTC method.

5.3.2 Random L_k and $P_a = 1$

In this subsection, we evaluate the scenario when the number of packets L_k sent by each device is random and follows a discrete uniform distribution as $L_k \sim \mathcal{U}[1,10]$, while we still maintain $P_a = 1$ such that $K = K_a$. Figures 33 and 34 show, respectively, the graphs of latency and network throughput. One can note that the superiority of the results achieved by the proposed methods over the results achieved by the iQmMTC and baseline methods are preserved. Indeed, the curves in Figures 33 and 34 present practically the same shapes of the ones in Figures 25 and 26, respectively, but with a little performance degradation due to the reduction in the average number of transmitted packets, which limits the learning capability of the Q-Learning algorithm in seeking less congested pilots.

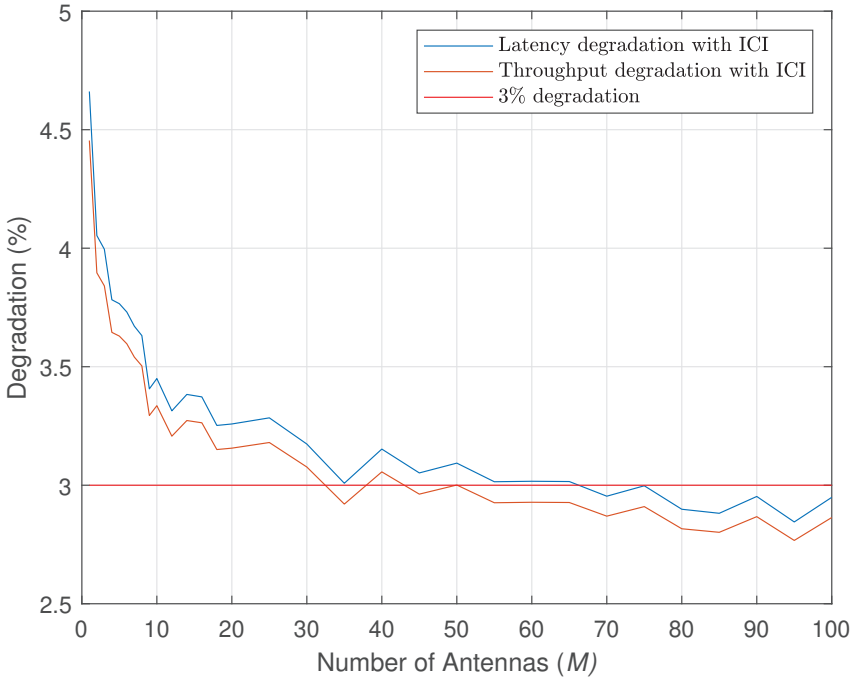


Figure 29 – Average network throughput and latency degradations $\times\, M$, for $L = 10$ packets, $K_a = 400$, and $K_{\text{ici}} = 400$.

5.3.3 Random L_k and $P_a = 0.1\%$

We consider, in this subsection, a random number of packets sent by each device following $L_k \sim \mathcal{U}[1,10]$, and a random number of devices being activated at each frame following a binomial distribution with activation probability of $P_a = 0.1\%$, such that $K_a \leq K_T$. The number of available RA pilots is also reduced to $\tau_p = 40$, in order to keep the simulation time not so long, and K_{ici} is kept fixed as 40. The reward in the QL framework of the mQmMTC GF RA protocol is calculated using (93), while assuming that $\mathbb{E}[L_k] = 5.5$ is known at the devices' side. The graphs presented in Figure 35 and 36 are generated with 64000 Monte-Carlo realizations.

Figure 35 presents the average per-user throughput and Figure 36 presents the average network throughput. In terms of both performance metrics, our proposed mQmMTC protocol remains quite close to that of the ideal cQmMTC protocol while always superior than that of Baseline and iQmMTC. While the per-user throughput of Baseline drops below 0.5 for $K_T \approx 2400$, this happens with $K_T \approx 2800$ for iQmMTC, with $K_T \approx 3100$ for mQmMTC, and with $K_T \approx 3400$ for the ideal cQmMTC. Similarly, the network throughput falling point in Figure 36 occurs with $K_T \approx 2300$ for the Baseline, with $K_T \approx 2400$ for iQmMTC, with $K_T \approx 2800$ for mQmMTC, and with $K_T \approx 3100$ for the ideal cQmMTC. These results corroborate the feasibility of the proposed mQmMTC GF RA protocol to address the challenges

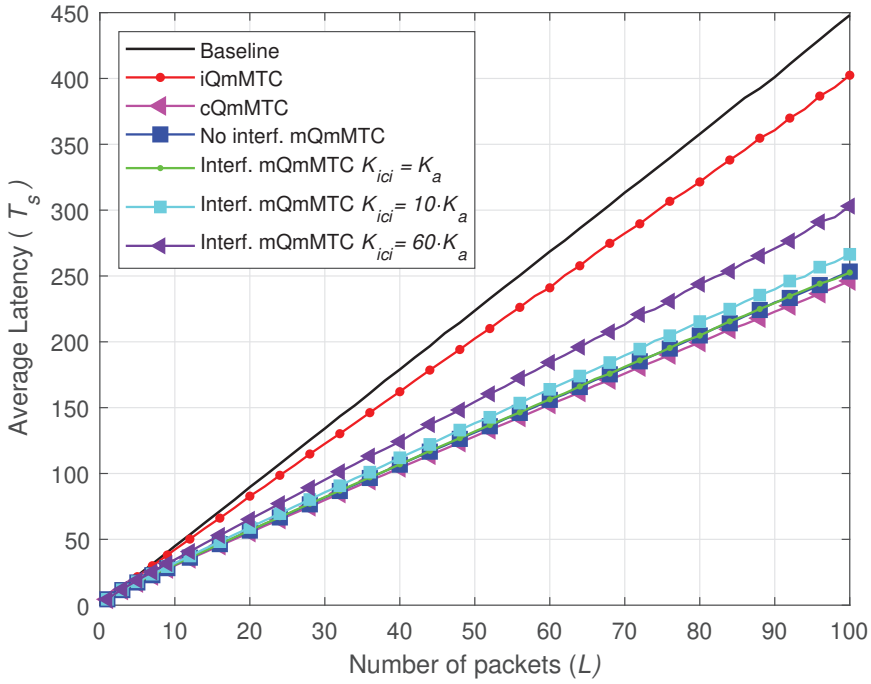


Figure 30 – Average latency $\times L$**, for** $K_a = 600$**,** $\tau_p = 400$**,** $K_{\text{ici}} = K_a$**,** $10 \cdot K_a$**,** $60 \cdot K_a$ **and** $M = 100$**.**

of massive machine-type communications in the framework of B5G systems. Nonetheless, the graph shows an inverted behavior between "No interf. mQmMTC" and "Interf. mQmMTC" chart lines when $K_T \in [2800; 3000]$ range. This fact needs further research to be explained. An initial hypothesis is that a higher number of Monte Carlo realizations is necessary to improve the accuracy of the graphs. Another hypothesis is that the model proposed in (92) is not accurate enough and should be replaced for a more accurate one. Some attempts to increase the number of Monte-Carlo repetitions were unsuccessful, firstly due to the simulation time which with 64 thousand Monte-Carlo repetitions is already 72 hours and secondly because of problems due to memory allocation. In the further research section 6.1, the development of a more efficient simulation computational environment is proposed, in which a faster algorithm and a more precise traffic predictor are considered.

5.4 CHAPTER CONCLUSIONS

In this chapter, we have applied the collaborative Q-Learning (QL) GF RA protocol to a practical massive MIMO scenario for pilot collision control, where we assumed realistic wireless propagation effects, such as multipath fading, shadowing, path loss, thermal noise and ICI. As the devices cannot know the exact number of pilot contenders without incurring excessive complexity and signaling overhead, our proposed approach takes advantage of the

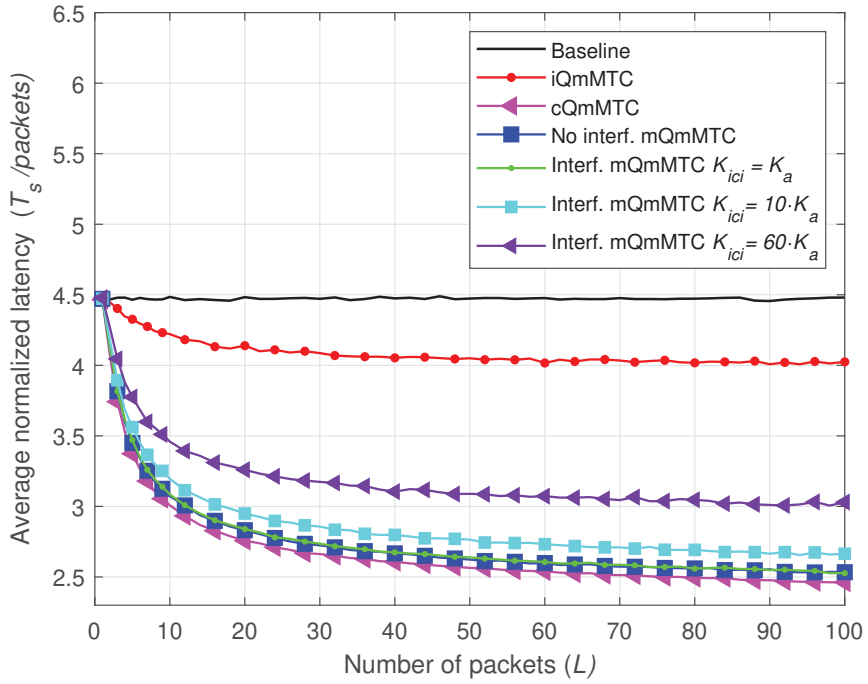


Figure 31 – Average normalized latency $\times L$, for $K_a = 600$, $\tau_p = 400$, $K_{\text{ici}} = K_a$, $10 \cdot K_a$, $60 \cdot K_a$ and $M = 100$.

massive number of BS antennas to allow the devices to compute the Q-learning rewards in a simplified way. We have also shown that our proposed approach is robust and does not require more than ≈ 30 BS antennas to produce significantly improved results, very close to the ideal cQmMTC protocol of (SHARMA; WANG, 2019). Furthermore, we have shown that the method also works for a small number of packets, such as 10, and when a number of packets and devices are random, following a uniform discrete distribution. This method has also shown robustness regarding the number of devices and levels of K_{ici}. Finally, we have shown that the number of activate devices can be random based on a binomial distribution.

Future works can decide on a SINR based criteria for decoding the payload in Equation (81) as well as investigating how many BS antennas would be adequate for an acceptable, successful decoding rate. Other works may include using a more sophisticated RLe-based approach, for instance, the *deep Q-learning* method (JIANG *et al.*, 2019; KUMAR *et al.*, 2021). It is also recommended the development of a more efficient computational algorithm for the simulations and a more precise traffic predictor, such the LSTM-based predictor (ABDELLAH; KOUCHERYAVY, 2020).

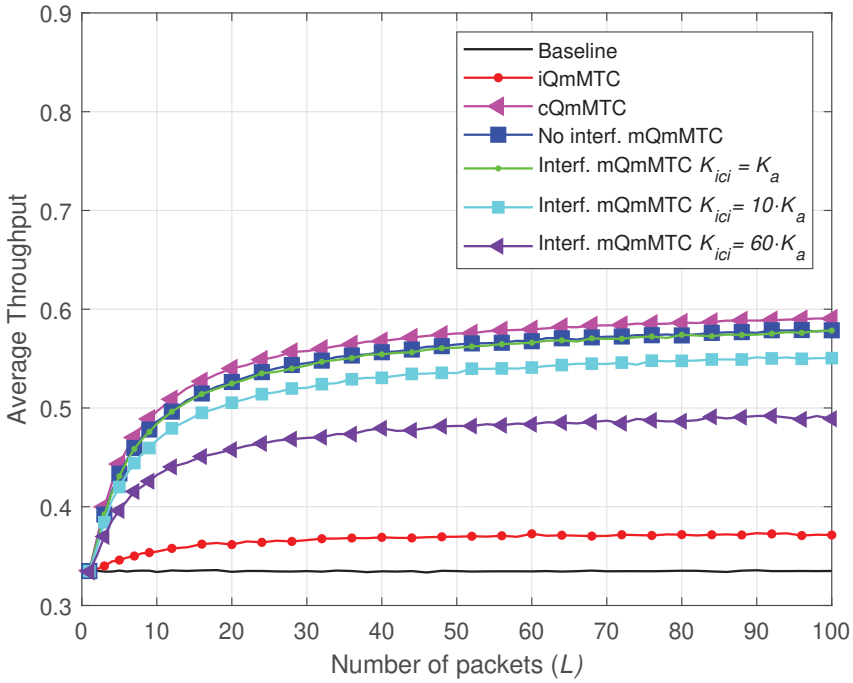


Figure 32 – Average network throughput $\times L$**, for** $K_a = 600$**,** $\tau_p = 400$**,** $K_{\text{ici}} = K_a$**,** $10 \cdot K_a$**,** $60 \cdot K_a$**, and** $M = 100$**.**

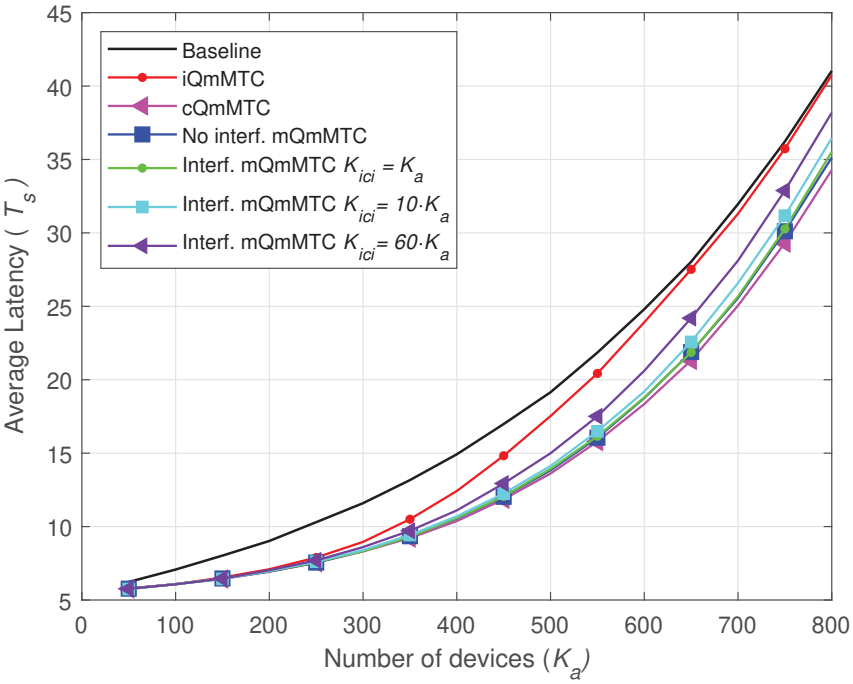


Figure 33 – Average latency $\times K_a$**, for** $L_k \sim \mathcal{U}$**(1,10),** $P_a = 1$**,** $\tau_p = 400$**,** $K_{\text{ici}} = K_a$**,** $10 \cdot K_a$**,** $60 \cdot K_a$**, and** $M = 100$**.**

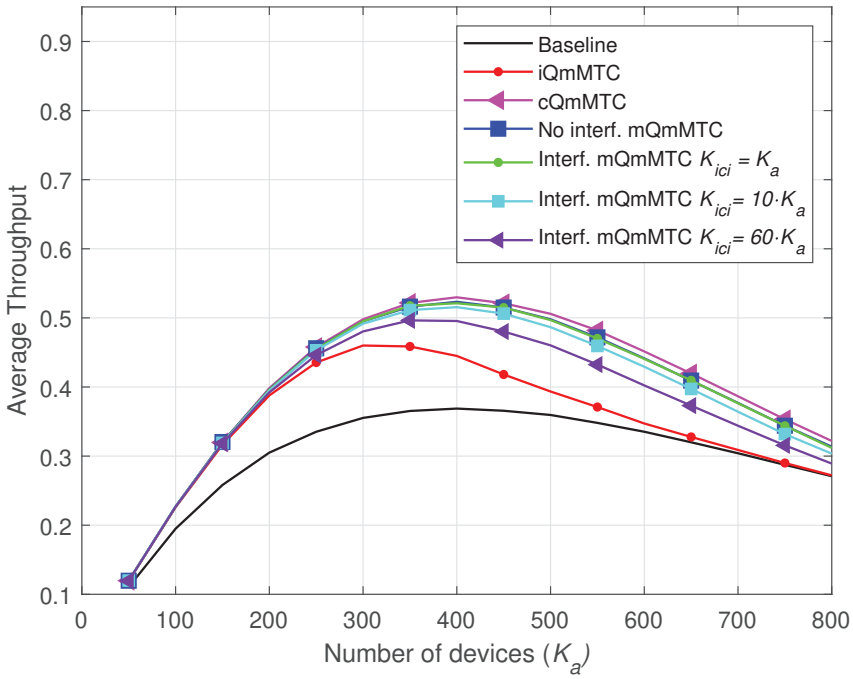


Figure 34 – Average network throughput $\times K_a$**, for** $L_k \sim \mathcal{U}$**(1,10),** $P_a = 1$**,** $\tau_p = 400$**,** $K_{\text{ici}} = K_a$**,** $10 \cdot K_a$**,** $60 \cdot K_a$**, and** $M = 100$**.**

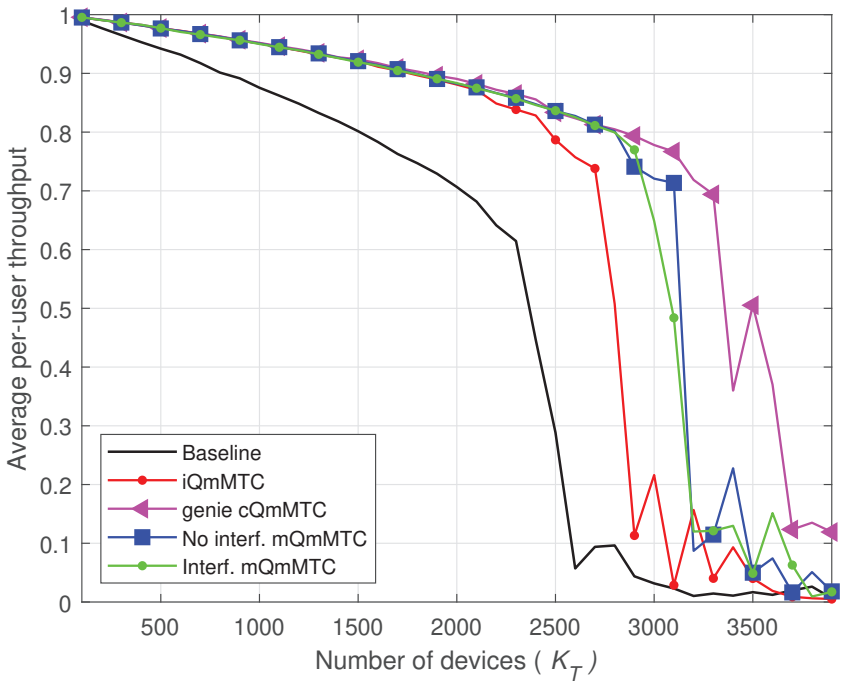


Figure 35 – Average per-user throughput $\times K_T$**, for** $L_k \sim \mathcal{U}$**(1,10),** $\tau_p = 40$**,** $K_{\text{ici}} = 40$**,** $P_a = 0.001$ **and** $M = 100$**.**

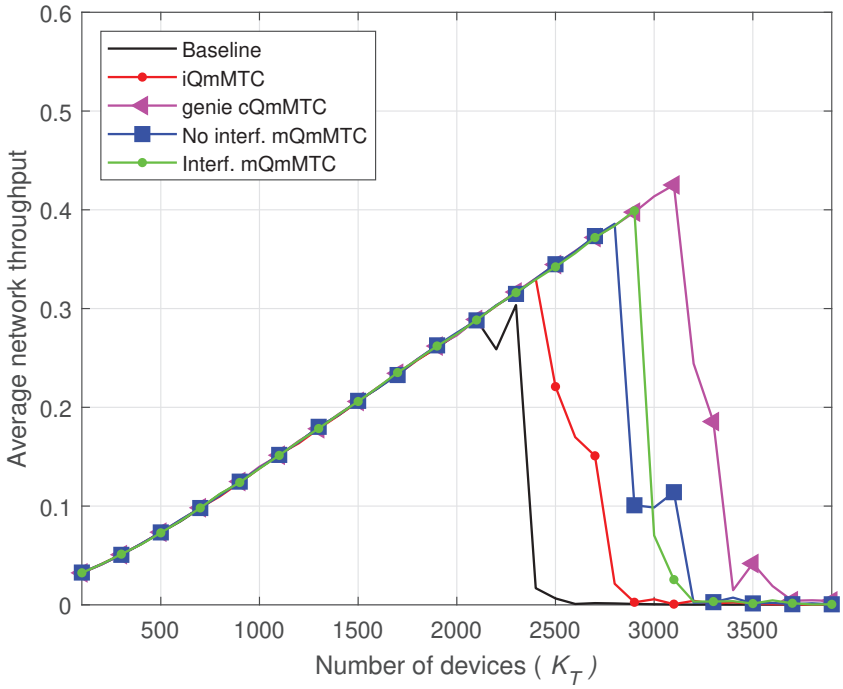


Figure 36 – Average network throughput $\times K_T$**, for** $L_k \sim \mathcal{U}$**(1,10),** $\tau_p = 40$**,** $K_{\text{ici}} = 40$**,** $P_a = 0.001$ **and** $M = 100$**.**

6 CONCLUSION

The increasing number of MTC devices represents a challenge when it comes to providing high connectivity and low latency in massive MIMO systems. Among many issues that may occur, there is *pilot collision*, which happens, in a massive MIMO system, when two or more devices in a cell send the same pilot to the BS for channel estimation. This problem represents a performance bottleneck in the system, impairing IoT and mMTC applications that usually require low latency. This work has investigated the use of statistical inference and machine learning based approaches for pilot collision control in massive MIMO systems. In Chapter 3 we propose a GB RA protocol for pilot collision resolution in massive MIMO systems, similar to the SUCRe protocol. However, our approach optimizes the decentralized decision making step of the SUCRe protocol with a BC classifier that indicates if the device is more likely to have the strongest signal among the contenders for particular pilot. Our proposed methodology trades off a slight increase in false-positive rates for a significant decrease in false-negative probabilities, resulting in significantly higher correct classification probabilities for the strongest user. This behavior is very beneficial in terms of connectivity performance, which has been significantly improved over the SUCRe protocol in both scenarios with and without ICI by our proposed method. The most significant performance gains were found in the overcrowded scenario, with more expected accessing devices than available RA pilots, i.e., $K_0 > \tau_p/P_a$, indicating the proposed method as a suitable strategy in massive MIMO systems with a very high number of devices. Furthermore, the results also show robustness regarding the number of antennas and edge SNR value.

In Chapter 4 a NN-based protocol is proposed as an alternative for the BC proposed in chapter Chapter 3. In the proposed methodology, a multilayer perceptron neural network is implemented as a classification tool to indicate if the device is the strongest contender. The results found indicate that the NN-based protocols present a slightly better performance than the BC classifier, in both scenarios with and without ICI, when $K_0 > 25000$, and only in scenarios with ICI, for a quantity of inactive UEs within a range of approximately 8000 UEs to approximately 12000 UEs. This second protocol also presents robustness concerning variation in the number of antennas, showing that the proposed machine-learning method can operate with a number of antennas other than it was trained.

In Chapter 5, a Q-Learning-based GF RA approach was introduced. This model was applied to a massive MIMO system, assuming realistic propagation effects, such as multipath

fading, shadowing, path loss, thermal noise and inter-cell interference. Because the devices cannot know the exact number of pilot candidates without incurring in excessive complexity and signaling overhead, our proposed approach makes use of the massive number of BS antennas to allow the devices to compute the Q-learning rewards in a simplified manner. We also demonstrated that our proposed approach is robust in terms of the number of packets to transmit, which can be as low as 10 or as random as a discrete uniform distribution, K_{ici} levels proportional to the number of active user in the center cell, and the number of active users in the center cell, which can be randomly activated as a binomial distribution. Our proposed method is also robust in terms of antenna number variation. Given that it does not require more than ≈ 30 antennas at the BS to produce significantly improved performance, which is very close to the ideal (*genie*) cQmMTC protocol of (SHARMA; WANG, 2019).

We can conclude that the three approaches, BC, NN and Q-Learning are useful tools of RA protocols in massive MIMO systems. In the next section some future works suggestions are provided.

6.1 FURTHER RESEARCH TOPICS

This section points out possible research idea that can be complementary to this work. Future research topics could include developing a new RA model for massive MIMO systems using a different ML or SI approach. Another option is to use a previously validated RA protocol in scenarios not previously studied. These potential future research topics could include one or more of the following ideas:

- Apply ULe based methods for either GB or GF RA algorithms. ULe algorithms are useful for clustering data in wide variety of applications. However, there is a lack of RA protocols that make use of ULe-based methods in the literature, which make their use appealing for a high-impact publication.

- Propose analytical methods for deriving the PDF of ϕ_k in chapter 3. A closed form expression for the PDF of ϕ_k allows for more precise and efficient computation and manipulation of the distribution and is more suitable for modeling complex relationships and dependencies between variables.

- Investigate the tendency of changing of threshold ϕ_k in chapter 3 for changes in different parameters such as P_a, P_r and K_{ici}. It is expected that changing the simulation parameters

of the threshold ϕ_k will also change. In relation to K_{ici}, there is little difference in performance without interference and with interference, which indicates a certain robustness to the presence and variation of ICI. Therefore, a variation of K_{ici} is not expected to modify the threshold ϕ_k too much. Regarding the activation probability P_a, if P_a assumes a value greater than the current one, a greater number of active users is also expected, and therefore, it is also expected that the threshold value ϕ_k^{th} will increase since with more active users the probability of the user being the strongest decreases. The same can be said about the probability P_r of a retransmission attempt. This analysis is important for providing more reliable and robust results.

- Propose different neural network topologies or architecture for the classification task in chapter 3. For example, implement the cost function binary cross entropy.

- Evaluate the behavior of the bias term ϵ_k in different usage scenarios.

- Increase the number K_{ici} proportionally to the number of UEs or devices. The purpose is to make the simulations more realistic as the number of interfering devices in neighboring cells are proportional to the number of devices in the center cell.

- Evaluate the performance of the NN proposed in chapter 4, considering different levels of transmission power and ICI.

- Propose a more realistic simulation scenario for IoT and mMTC applications, *e.g.*, a smart factory plant. Since the IoT has become a commercial trend, to aim for always more realistic models has become of paramount importance.

- Propose a coefficient to evaluate the performance of the RLe method in 5 proportional to the levels of K_{ici}, the number of devices and the angular coefficient κ_k of the curve in Figure 30 (Latency $\times$ L). One option is the following.

$$\alpha_k = \kappa_k \cdot \frac{K_{\text{ici}}}{K_a} \tag{94}$$

- Integrate the IRS and the mmWave technologies to the proposed methods and evaluate their performances.

- Discuss the trade-off between time division, which results in less interference but less efficient usage of communication resources, and and temporal superposition, already

applied to the mQmMTC protocol, which leads to more interference that could be mitigated when applying massive MIMO.

- Analyze the SINR behaviour of the protocol proposed in Chapter 5 and investigate how many antennas would be necessary to achieve an acceptable SINR value according to Equation 85 converted to dBs with $\gamma_{dBk}^{\text{ul}} = 10\log_{10}(\gamma_k^{\text{ul}})$. Some previous results are shown in Figures 37 and 38. In Figure 37 is shown the average SINR, for cQmMTC and mQmMTC for the cases with and without ICI, for M varying from 0 to 2000.

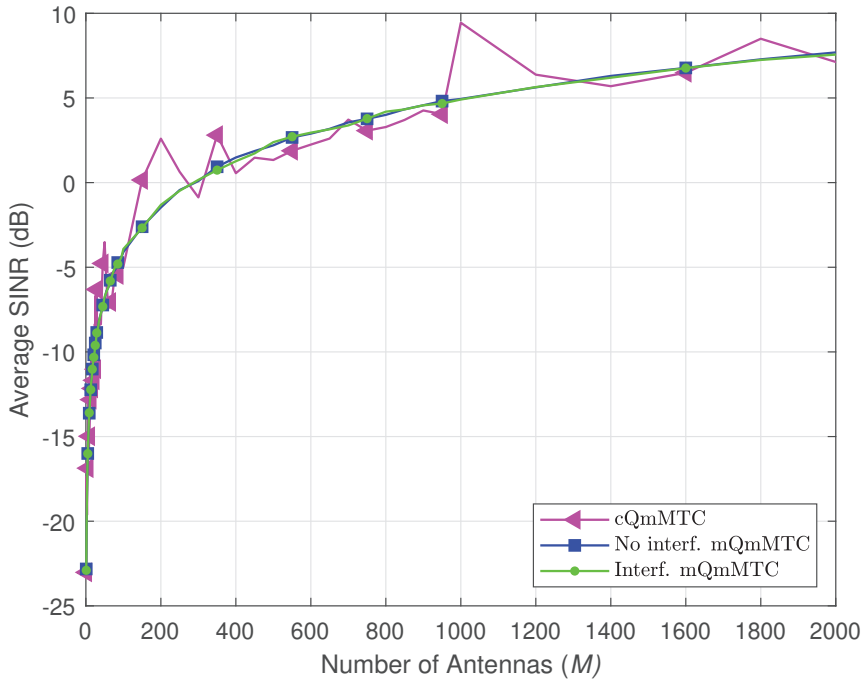


Figure 37 – Average SINR $\times$ M (for $L = 10$ packets, $K_a = 100$, $\tau_p = K_{ici} = 100$)

Figure 38 shows the Complementary Cumulative distribution function (CCDF) for $M = 100$, $M = 500$, $M = 1000$ e $M = 2000$, $L = 10$ packets, $K = 100$, and $\tau_p = K_{ici} = 100$. It indicates at each SINR level ranging from -20 to 15 dB the probability that the actual SINR of the wireless channel will be above that level.

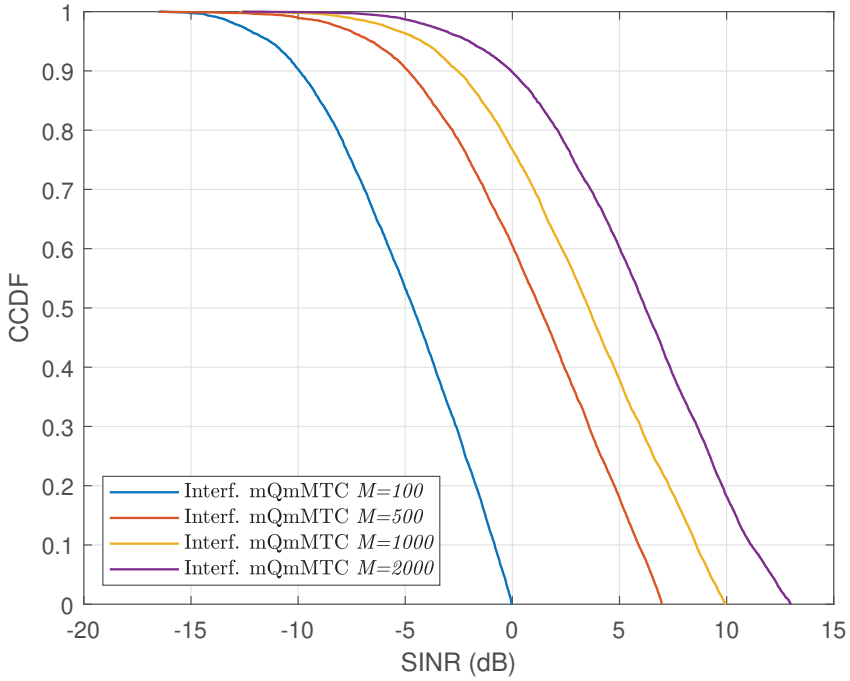


Figure 38 – (CCDF) of uplink SINR in (dB) for $M = 100$, $M = 500$, $M = 1000$ e $M = 2000$ (for $L = 10$ packets, $K_a = 100$, $\tau_p = K_{ici} = 100$)

.

REFERENCES

3GPP. *Feasibility Study on New Services and Markets Technology Enablers*. Valbonne, Fr, 2016.

3GPP. *Spatial channel model for Multiple Input Multiple Output (MIMO) simulations*. Valbonne, Fr, 2018. Version 15.0.0.

5G PPP. *PPP CONTRACT*. 2013. Available at: <https://5g-ppp.eu/contract/>. Accessed on: 06.07.2022.

ABDELLAH, Ali R.; KOUCHERYAVY, Andrey. Deep Learning with Long Short-Term Memory for IoT Traffic Prediction. *In*: GALININA, Olga; ANDREEV, Sergey; BALANDIN, Sergey; KOUCHERYAVY, Yevgeni (Ed.). **Internet of Things, Smart Spaces, and Next Generation Networks and Systems**. Cham: Springer International Publishing, 2020. p. 267–280.

AFSHAR, Abolghasem; VAKILI, Vahid Tabataba. Spatial Based Pilot Allocation (SBPA) in Crowded Massive MIMO Systems. **Wireless Personal Communications**, v. 119, n. 1, p. 239–257, Jul 2021. ISSN 1572-834X.

AL-FALAHY, Naser; ALANI, Omar Y. Technologies for 5G Networks: Challenges and Opportunities. **IT Professional**, v. 19, n. 1, p. 12–20, 2017.

ALEXANDERSON, E. F. W. Transoceanic radio communication. **Transactions of the American Institute of Electrical Engineers**, XXXVIII, n. 2, p. 1269–1285, 1919.

BAI, Lin; LIU, Jiexun; YU, Quan; CHOI, Jinho; ZHANG, Wei. A collision resolution protocol for random access in massive mimo. **IEEE Journal on Selected Areas in Communications**, v. 39, n. 3, p. 686–699, 2021.

BELLO, Lawal Mohammed; MITCHELL, Paul; GRACE, David. Application of Q-Learning for RACH Access to Support M2M Traffic over a Cellular Network. *In*: **European Wireless 2014; 20th European Wireless Conference**. Berlin, Germany: VDE VERLAG, 2014. p. 1–6. ISBN 9783800736218.

BISHOP, Christopher M. **Pattern Recognition and Machine Learning (Information Science and Statistics)**. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738.

BJÖRNSON, Emil. **Six Differences Between MU-MIMO And massive MIMO**. 2017. Available at: https://ma-mimo.ellintech.se/2017/10/17/six-differences-between-mu-mimo-and-massive-mimo/. Accessed on: 05.07.2022.

BJÖRNSON, Emil; de Carvalho, E.; Sørensen, J. H.; Larsson, E. G.; Popovski, P. A Random Access Protocol for Pilot Allocation in Crowded Massive MIMO Systems. **IEEE Transactions on Wireless Communications**, v. 16, n. 4, p. 2220–2234, April 2017.

BJÖRNSON, Emil; HOYDIS, Jakob; SANGUINETTI, Luca. Massive mimo networks: Spectral, energy, and hardware efficiency. **Foundations and Trends® in Signal Processing**, v. 11, n. 3-4, p. 154–655, 2017. ISSN 1932-8346.

BJÖRNSON, Emil; LARSSON, Erik G.; MARZETTA, Thomas L. Massive MIMO: Ten Myths and One Critical Question. **IEEE Communications Magazine**, v. 54, n. 2, p. 114–123, 2016.

BJÖRNSON, Emil; SANGUINETTI, Luca; WYMEERSCH, Henk; HOYDIS, Jakob; MARZETTA, Thomas L. Massive MIMO is a Reality - What is Next?: Five Promising Research Directions for Antenna Arrays. **Digital Signal Processing**, v. 94, p. 3–20, 2019. ISSN 1051-2004.

BUENO, Felipe Augusto Dutra; YAMAMURA, Cézar Fumio; GOEDTEL, Alessandro; FILHO, José Carlos Marinello. A Random Access Protocol for Crowded Massive MIMO Systems Based on a Bayesian Classifier. **IEEE Wireless Communications Letters**, v. 11, n. 11, p. 2455–2459, 2022.

BZDOK, Danilo; ALTMAN, Naomi; KRZYWINSKI, Martin. Statistics versus machine learning. **Nature Methods**, v. 15, n. 4, p. 233–234, Apr 2018. ISSN 1548-7105.

CASINI, Enrico; GAUDENZI, Riccardo De; HERRERO, Oscar Del Rio. Contention Resolution Diversity Slotted ALOHA (CRDSA): An Enhanced Random Access Scheme for Satellite Access Packet Networks. **IEEE Transactions on Wireless Communications**, v. 6, n. 4, p. 1408–1419, 2007.

CHEN, Xiaoming; NG, Derrick Wing Kwan; YU, Wei; LARSSON, Erik G.; AL-DHAHIR, Naofal; SCHOBER, Robert. Massive Access for 5G and Beyond. **IEEE Journal on Selected Areas in Communications**, v. 39, n. 3, p. 615–637, 2021.

CHERGUI, Hatim; BLANCO, Luis; VERIKOUKIS, Christos. Statistical Federated Learning for Beyond 5G SLA-Constrained RAN Slicing. **IEEE Transactions on Wireless Communications**, v. 21, n. 3, p. 2066–2076, 2022.

CHUN, Chang-Jae; KANG, Jae-Mo; KIM, Il-Min. Deep Learning-Based Channel Estimation for Massive MIMO Systems. **IEEE Wireless Communications Letters**, v. 8, n. 4, p. 1228–1231, 2019.

CISCO. **Ericsson Mobility Report**. 2020. Available at: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. Accessed on: 22.06.2022.

CIVIL HOUSE. **Governo Federal oficializa a concessão do 5G** . 2021. Available at: https://www.gov.br/casacivil/pt-br/assuntos/noticias/2021/dezembro/governo-federal-oficializa-a-concessao-do-5g. Accessed on: 28.06.2022.

CRÂŞMARIU, Victor-Florin; ARVINTE, Marius-Octavian; ENESCU, Andrei-Alexandru; CIOCHINA, Silviu. Waterfilling power allocation applied in block-diagonalization multi-user mimo precoding technique. *In*: **2016 12th IEEE International Symposium on Electronics and Telecommunications (ISETC)**. *[S.l.: s.n.]*, 2016. p. 91–94.

CUI, Jingjing; DING, Zhiguo; FAN, Pingzhi; AL-DHAHIR, Naofal. Unsupervised Machine Learning-Based User Clustering in Millimeter-Wave-NOMA Systems. **IEEE Transactions on Wireless Communications**, v. 17, n. 11, p. 7425–7440, 2018.

DING, Jie; CHOI, Jinho. Preamble-Data Superposition Random Access in Massive MIMO Systems. **IEEE Wireless Communications Letters**, v. 9, n. 6, p. 906–910, 2020.

DING, Jie; QU, Daiming; JIANG, Hao; JIANG, Tao. Success Probability of Grant-Free Random Access With Massive MIMO. **IEEE Internet of Things Journal**, v. 6, n. 1, p. 506–516, 2019.

DUDA, Richard O.; HART, Peter E.; STORK, David G. **Pattern Classification (2nd Edition)**. 2. ed. *[S.l.]*: Wiley-Interscience, 2000. ISBN 0471056693.

ERICSSON. **Ericsson Mobility Report**. 2022. Available at: https://www.ericsson.com/49d3a0/assets/local/reports-papers/mobility-report/documents/2022/ericsson-mobility-report-june-2022.pdf. Accessed on: 22.06.2022.

FARSAD, Nariman; GOLDSMITH, Andrea. **Detection Algorithms for Communication Systems Using Deep Learning**. arXiv, 2017. Available at: https://arxiv.org/abs/1705.08044.

FEHRENBACH, Thomas; DATTA, Rohit; GöKTEPE, Bariş; WIRTH, Thomas; HELLGE, Cornelius. URLLC Services in 5G Low Latency Enhancements for LTE. *In*: **2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)**. New York, NY: IEEE, 2018. p. 1–6.

GAO, Xiang; EDFORS, Ove; RUSEK, Fredrik; TUFVESSON, Fredrik. Massive MIMO Performance Evaluation Based on Measured Propagation Data. **IEEE Transactions on Wireless Communications**, v. 14, n. 7, p. 3899–3911, 2015.

GOLDSMITH, Andrea. **Wireless Communications**. Cambridge, UK: Cambridge University Press, 2005.

GóMEZ-ANDRADES, Ana; MUñOZ, Pablo; SERRANO, Inmaculada; BARCO, Raquel. Automatic root cause analysis for lte networks based on unsupervised techniques. **IEEE Transactions on Vehicular Technology**, v. 65, n. 4, p. 2369–2386, 2016.

HAN, Huimei; GUO, Xudong; LI, Ying. A High Throughput Pilot Allocation for M2M Communication in Crowded Massive MIMO Systems. **IEEE Transactions on Vehicular Technology**, v. 66, n. 10, p. 9572–9576, 2017a.

HAN, H.; Li, Y.; Guo, X. A Graph-Based Random Access Protocol for Crowded Massive MIMO Systems. **IEEE Transactions on Wireless Communications**, v. 16, n. 11, p. 7348–7361, 2017b.

HAN, Huimei; LI, Ying; ZHAI, Wenchao; QIAN, Liping. A Grant-Free Random Access Scheme for M2M Communication in Massive MIMO Systems. **IEEE Internet of Things Journal**, v. 7, n. 4, p. 3602–3613, 2020.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. New York, NY: Springer, 2009. ISSN 0172-7397. ISBN 978-0-387-84857-0.

HOFF, Peter D. **A first course in Bayesian statistical methods**. New York, NY: Springer, 2009. v. 580.

HUANG, Hongji; SONG, Yiwei; YANG, Jie; GUI, Guan; ADACHI, Fumiyuki. Deep-learning-based millimeter-wave massive mimo for hybrid precoding. **IEEE Transactions on Vehicular Technology**, v. 68, n. 3, p. 3027–3032, 2019.

HUANG, Hongji; YANG, Jie; HUANG, Hao; SONG, Yiwei; GUI, Guan. Deep learning for super-resolution channel estimation and doa estimation based massive mimo system. **IEEE Transactions on Vehicular Technology**, v. 67, n. 9, p. 8549–8560, 2018.

IBM. **What is machine learning?** . 2023. Available at: https://www.ibm.com/topics/machine-learning. Accessed on: 07.01.2022.

JIANG, Nan; DENG, Yansha; SIMEONE, Osvaldo; NALLANATHAN, Arumugam. Cooperative Deep Reinforcement Learning for Multiple-group NB-IoT Networks Optimization. *In*: **ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. *[S.l.: s.n.]*, 2019. p. 8424–8428.

KOLODZIEJ, Kenneth E.; COOKSON, Aidan U.; PERRY, Bradley T. Machine Learning for Accelerated IBFD Tuning in 5G Flexible Duplex Networks. *In*: **2020 IEEE/MTT-S International Microwave Symposium (IMS)**. New York, NY: IEEE, 2020. p. 691–694.

KUMAR, Abhishek; VERMA, Gunjan; RAO, Chirag; SWAMI, Ananthram; SEGARRA, Santiago. Adaptive Contention Window Design Using Deep Q-Learning. *In*: **ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. *[S.l.: s.n.]*, 2021. p. 4950–4954.

LARSSON, Erik G.; EDFORS, Ove; TUFVESSON, Fredrik; MARZETTA, Thomas L. Massive MIMO for next generation wireless systems. **IEEE Communications Magazine**, v. 52, n. 2, p. 186–195, 2014.

LIU, Guanying; LIU, An; LIAN, Lixiang; LAU, Vincent; ZHAO, Min-Jian. Sparse Bayesian Inference Based Direct Localization for Massive MIMO. *In*: **2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)**. New York, NY: IEEE, 2019. p. 1–5.

LIVA, Gianluigi. Graph-Based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA. **IEEE Transactions on Communications**, v. 59, n. 2, p. 477–487, 2011.

LU, Lu; LI, Geoffrey Ye; SWINDLEHURST, A. Lee; ASHIKHMIN, Alexei; ZHANG, Rui. An Overview of Massive MIMO: Benefits and Challenges. **IEEE Journal of Selected Topics in Signal Processing**, v. 8, n. 5, p. 742–758, 2014.

MAHMOOD, Nurul Huda; BÖCKER, Stefan; MUNARI, Andrea; CLAZZER, Federico; MOERMAN, Ingrid; MIKHAYLOV, Konstantin; LOPEZ, Onel; PARK, Ok-Sun; MERCIER, Eric; BARTZ, Hannes; JäNTTI, Riku; PRAGADA, Ravikumar; MA, Yihua; ANNANPERÄ, Elina; WIETFELD, Christian; ANDRAUD, Martin; LIVA, Gianluigi; CHEN, Yan; GARRO, Eduardo; BURKHARDT, Frank; ALVES, Hirley; LIU, Chen-Feng; SADI, Yalcin; DORE, Jean-Baptiste; KIM, Eunah; SHIN, JaeSheung; PARK, Gi-Yoon; KIM, Seok-Ki; YOON, Chanho; ANWAR, Khoirul; SEPPÄNEN, Pertti. **White Paper on Critical and Massive Machine Type Communication Towards 6G**. arXiv, 2020. 3 p. Available at: https://arxiv.org/abs/2004.14146.

MARINELLO, J. C.; ABRÃO, T. Collision Resolution Protocol via Soft Decision Retransmission Criterion. **IEEE Transactions on Vehicular Technology**, v. 68, n. 4, p. 4094–4097, 2019.

MARINELLO, J. C.; ABRÃO, T.; Souza, R. D.; de Carvalho, E.; Popovski, P. Achieving Fair Random Access Performance in Massive MIMO Crowded Machine-Type Networks. **IEEE Wireless Communications Letters**, v. 9, n. 4, p. 503–507, 2020.

MARINELLO, José Carlos; BRANTE, Glauber; SOUZA, Richard Demo; ABRÃO, Taufik. Exploring the Non-Overlapping Visibility Regions in XL-MIMO Random Access and Scheduling. **IEEE Transactions on Wireless Communications**, p. 1–1, 2022.

MARINELLO, José Carlos; PANAZIO, Cristiano; ABRÃO, Taufik. Uplink performance of single-carrier receiver in massive mimo with pilot contamination. **IEEE Access**, v. 5, p. 8669–8681, 2017.

MARZETTA, Thomas L. Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas. **IEEE Transactions on Wireless Communications**, v. 9, n. 11, p. 3590–3600, 2010.

MARZETTA, Thomas L. Massive MIMO: An Introduction. **Bell Labs Technical Journal**, v. 20, p. 11–22, 2015.

MARZETTA, Thomas L.; LARSSON, Erik G.; YANG, Hong; NGO, Hien Quoc. **Fundamentals of Massive MIMO**. Cambridge, UK: Cambridge University Press, 2016. 21 p.

Ministry of Communication. **Leilão do 5G deve movimentar R$ 169 bilhões em investimentos** . 2021. Available at: https://www.gov.br/mcom/pt-br/noticias/2021/outubro/leilao-do-5g-deve-movimentar-r-169-bilhoes-em-investimentos. Accessed on: 28.06.2022.

MUSTAFA, Raza Ul; MOURA, David; ROTHENBERG, Christian Esteve. Machine Learning Approach to Estimate Video QoE of Encrypted DASH Traffic in 5G Networks. *In*: **2021 IEEE Statistical Signal Processing Workshop (SSP)**. New York, NY: IEEE, 2021. p. 586–589.

NGMN. **5G white paper**. 2015. Available at: https://ngmn.org/wp-content/uploads/NGMN_5G_White_Paper_V1_0.pdf. Accessed on: 30.06.2022.

NGO, Hien Quoc; LARSSON, Erik G. No Downlink Pilots Are Needed in TDD Massive MIMO. **IEEE Transactions on Wireless Communications**, v. 16, n. 5, p. 2921–2935, 2017.

NISHIMURA, Otávio Seidi; MARINELLO, José Carlos; ABRÃO, Taufik. A Grant-Based Random Access Protocol in Extra-Large Massive MIMO System. **IEEE Communications Letters**, v. 24, n. 11, p. 2478–2482, 2020.

NOMEIR, Mohamed W.; GADALLAH, Yasser; SEDDIK, Karim G. Uplink Scheduling for Mixed Grant-Based eMBB and Grant-Free URLLC Traffic in 5G Networks. *In*: **2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)**. New York, NY: IEEE, 2021. p. 187–192.

OROZA, Carlos A.; GIRALDO, Jairo A.; PARVANIA, Masood; WATTEYNE, Thomas. Wireless-Sensor Network Topology Optimization in Complex Terrain: A Bayesian Approach. **IEEE Internet of Things Journal**, v. 8, n. 24, p. 17429–17435, 2021.

PEREIRA, Hebert Douglas; BRANTE, Glauber; FARHAT, Jamil; SOUZA, Richard Demo; FILHO, José Carlos Marinello; ABRÃO, Taufik. On the Sum-Rate of Contention Resolution in Massive MIMO With NOMA. **IEEE Access**, v. 9, p. 24965–24974, 2021.

REKKAS, Vasileios P; SOTIROUDIS, Sotirios; SARIGIANNIDIS, Panagiotis; WAN, Shaohua; KARAGIANNIDIS, George K; GOUDOS, Sotirios K. Machine learning in beyond 5g/6g networks—state-of-the-art and future trends. **Electronics**, MDPI, v. 10, n. 22, p. 2786, 2021.

REN, Jie; WANG, Zulin; XU, Mai; FANG, Fang; DING, Zhiguo. An EM-Based User Clustering Method in Non-Orthogonal Multiple Access. **IEEE Transactions on Communications**, v. 67, n. 12, p. 8422–8434, 2019.

REZAZADEH, Farhad; CHERGUI, Hatim; BLANCO, Luis; ALONSO, Luis; VERIKOUKIS, Christos. A Collaborative Statistical Actor-Critic Learning Approach for 6G Network Slicing Control. *In*: **2021 IEEE Global Communications Conference (GLOBECOM)**. New York, NY: IEEE, 2021. p. 1–6.

RUSEK, Fredrik; PERSSON, Daniel; LAU, Buon Kiong; LARSSON, Erik G.; MARZETTA, Thomas L.; EDFORS, Ove; TUFVESSON, Fredrik. Scaling up mimo: Opportunities and challenges with very large arrays. **IEEE Signal Processing Magazine**, v. 30, n. 1, p. 40–60, 2013.

SANTOS, Herman Lucas dos; MARINELLO, José Carlos; PANAZIO, Cristiano Magalhaes; ABRÃO, Taufik. Machine learning-aided pilot and power allocation in multi-cellular massive MIMO networks. **Physical Communication**, v. 52, p. 101646, 2022. ISSN 1874-4907.

SENEL, Kamil; LARSSON, Erik G. Grant-free massive mtc-enabled massive mimo: A compressive sensing approach. **IEEE Transactions on Communications**, v. 66, n. 12, p. 6164–6175, 2018.

SEO, Jun-Bae; JUNG, Bang Chul; JIN, Hu. Modeling and Online Adaptation of ALOHA for Low-Power Wide-Area Networks (LPWANs). **IEEE Internet of Things Journal**, v. 8, n. 20, p. 15608–15619, 2021.

SHAHAB, Muhammad Basit; ABBAS, Rana; SHIRVANIMOGHADDAM, Mahyar; JOHNSON, Sarah J. Grant-free non-orthogonal multiple access for iot: A survey. **IEEE Communications Surveys Tutorials**, v. 22, n. 3, p. 1805–1838, 2020.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding Machine Learning: From Theory to Algorithms**. USA: Cambridge University Press, 2014. ISBN 1107057132.

SHARMA, Shree Krishna; WANG, Xianbin. Collaborative Distributed Q-Learning for RACH Congestion Minimization in Cellular IoT Networks. **IEEE Communications Letters**, v. 23, n. 4, p. 600–603, 2019.

SILVA, I.N. da; SPATTI, D.H.; FLAUZINO, R.A.; LIBONI, L.H.B.; ALVES, S.F. dos Reis. **Artificial Neural Networks: A Practical Course**. *[S.l.]*: Springer International Publishing, 2016. ISBN 9783319431628.

SJ, Gerasenko; JOSHI, Abhijit; RAYAPROLU, Srinivas; PONNAVAIKKO, Kovendhan; AGRAWAL, Dharma. Beacon signals: What, why, how, and where? **Computer**, v. 34, p. 108 – 110, 11 2001.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement Learning: An Introduction**. 2. ed. *[S.l.]*: The MIT Press, 2018.

TAO, Yiwen; LI, Bin; ZHAO, Chenglin. Channel Detection Under Impulsive Noise and Fading Environments for Smart Grid. **IEEE Internet of Things Journal**, v. 8, n. 20, p. 15407–15421, 2021.

THOOTA, Sai Subramanyam; MURTHY, Chandra R. Variational Bayesian Inference based Soft-Symbol Decoding for Uplink Massive MIMO Systems with Low Resolution ADCs. *In*: **2019 53rd Asilomar Conference on Signals, Systems, and Computers**. New York, NY: IEEE, 2019. p. 2180–2184.

WANG, Fanggang; MA, Guoyu. **Massive Machine Type Communications: Multiple Access Schemes**. First. Cham, Switzerland: Springer Cham, 2019. 1 p. ISSN 2191-8112.

WEN, Chao-Kai; SHIH, Wan-Ting; JIN, Shi. Deep Learning for Massive MIMO CSI Feedback. **IEEE Wireless Communications Letters**, v. 7, n. 5, p. 748–751, 2018.

WU, Qingqing; ZHANG, Rui. Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming. **IEEE Transactions on Wireless Communications**, v. 18, n. 11, p. 5394–5409, 2019.

YU, Hanxiao; FEI, Zesong; ZHENG, Zhong; YE, Neng. Finite-Alphabet Signature Design for Grant-Free NOMA: A Quantized Deep Learning Approach. **IEEE Transactions on Vehicular Technology**, v. 69, n. 10, p. 10975–10987, 2020.

ZHOU, Yibo; FADLULLAH, Zubair Md.; MAO, Bomin; KATO, Nei. A Deep-Learning-Based Radio Resource Assignment Technique for 5G Ultra Dense Networks. **IEEE Network**, v. 32, n. 6, p. 28–34, 2018.

APPENDIX

APPENDIX A – A RANDOM ACCESS PROTOCOL FOR CROWDED MASSIVE MIMO SYSTEMS BASED ON A BAYESIAN CLASSIFIER

A Random Access Protocol for Crowded Massive MIMO Systems Based on a Bayesian Classifier

Felipe Augusto Dutra Bueno, Cézar Fumio Yamamura, Alessandro Goedtel, and José Carlos Marinello Filho

***Abstract*—The strongest user collision resolution (SUCRe) protocol is an efficient grant-based random access (RA) solution to provide connectivity for large numbers of user equipments (UEs), leveraging massive MIMO propagation features, typically available in fifth-generation and beyond networks. In this letter, we propose to replace the retransmission rule of SUCRe protocol by a Bayesian classifier for identifying the strongest user, aiming to resolve the collisions in a decentralized way, at the UEs' side. As an offline training stage, we first conduct a statistical learning procedure to obtain the density estimations of the UEs' decision variables in the SUCRe protocol, both in the cases of the UE being the strongest contender or not. Then, following the maximum *a posteriori* decision criterion, we determine how the UE can decide if it is the strongest contender (retransmitting the chosen RA pilot) or not (staying idle and trying again later). The numerical results show that our proposed method achieves significant connectivity performance improvements compared with other protocols, without requiring any additional complexity or overhead.**

***Index Terms*—Random access, grant-based protocols, massive MIMO, statistical learning, machine learning.**

I. Introduction

THE NUMBER of connected devices is ever-increasing in fifth-generation (5G) and beyond (B5G) networks [1], while the available time and frequency communication resources remain scarce. This scarcity of resources gives rise to a connectivity performance bottleneck in such networks since pilot collisions during the random access (RA) stage for user equipments (UEs) become more frequent, especially critical for UEs requiring reliable and low latency communications. Therefore, the design and implementation of powerful and effective RA protocols are very important for the development and consolidation of B5G networks [1].

Several solutions have been proposed to mitigate the problem of pilot collisions. The strongest-user collision resolution (SUCRe) protocol proposed in [2] is a 4-steps grant-based procedure whose main idea is to allow only the strongest contender to access the network resources. In the first step, the UEs which want to become active transmit in the uplink (UL) an RA pilot sequence randomly chosen from a common pilots' set. Then, in step 2, the base station (BS) responds with a precoded pilot response in the downlink (DL), with the precoding evaluated from the channel estimates obtained from each received RA pilot. By receiving this signal, the UEs can estimate information about the sum of the signal strengths of all UEs contending for the same RA pilot. Based on this information, each contender decides, in a decentralized way, whether or not to retransmit its pilot signal in the UL of step 3. The retransmission occurs only when the UE signal strength corresponds to more than 50% of the total sum of the signal strengths of all UEs contending for a given RA pilot signal. Finally, in step 4, the BS allocates dedicated communication resources to the UEs who did not collide in step 3. The SUCRe protocol is very effective and can resolve 90% of collisions [2]. However, the protocol does not resolve pilot collisions in which the strongest UE has a signal strength lower than 50% of the sum of the signal strengths of contending UEs, characterizing false-negatives.

In [3] and [4], collision resolution schemes are proposed in which the UEs receive from the BS not only the precoded DL response to estimate the sum of the signal strengths of all contenders, but also the information on which pilots remain idle after the first RA round, aiming to admit as many UEs as possible to the network. The results of both studies demonstrate superior performance concerning the SUCRe method. However, the extra overhead required to inform the idle pilots needs to be taken into account for a fairer comparison, which increases the signaling overhead in the RA stage, increases latency and penalizes spectral efficiency.

In [5], a soft decision retransmission rule is proposed to improve the performance of the SUCRe protocol. This rule is based on the probability of the UE being the strongest contender for its chosen pilot. The obtained results are superior to the original SUCRe protocol, but requiring additional information known at the UEs' side. In [6], the access class barrier with power control (ACBPC) RA protocol is proposed. The results indicate a relative connectivity performance gain in comparison with the SUCRe protocol, while providing homogeneous access possibilities to the UEs, independent of their distances to the BS. Other works, such as [7], [8] and [9], also present different alternatives of collision resolution protocols. However, [7] and [8] introduce extra complexity and overhead to the RA stage, while in [9] the SUCRe protocol is adapted for extra-large MIMO (XL-MIMO) systems, taking advantage of the spatial non-stationarities typically available in such scenarios.

The Bayesian method is a classical inference statistical tool that compares the weighted probability density functions (PDFs) of numerous classes and costs associated with them, and then selects the class having the highest odd of being true [10]. Although being a simple method, it can achieve high accuracy levels when coupled with kernel density estimation under statistical learning frameworks. It has been recently applied to solve problems like backlogged devices number estimation in low-power wide-area networks [11], channel detection in smart grid [12], and others.

Manuscript received 15 July 2022; revised 23 August 2022; accepted 10 September 2022. Date of publication 15 September 2022; date of current version 9 November 2022. The associate editor coordinating the review of this article and approving it for publication was S. Kafaie. *(Corresponding author: José Carlos Marinello Filho.)*

The authors are with the Electrical Engineering Department, Federal University of Technology—Paraná, Cornélio Procópio 86300-000, Brazil (e-mail: faugustobueno@gmail.com; cezaryamamura@gmail.com; agoedtel@utfpr.edu.br; jcmarinello@utfpr.edu.br).

Digital Object Identifier 10.1109/LWC.2022.3207069

This letter proposes to replace the retransmission rule of the original SUCRe protocol by a Bayesian classifier (BC) to indicate if the UE is the strongest contender or not, resolving thus the pilot collision in a decentralized way. The main contributions are as follow: **i)** We characterize the RA in crowded massive MIMO systems employing the SUCRe framework as a classification problem to the UEs decide whether they are the strongest contender for the chosen pilot or not, and propose a solution aiming to maximize the probability of correct classification by the UEs given their observations; **ii)** As the required PDFs for the BC computation are difficult to obtain analytically, we first carry out an offline statistical learning procedure to empirically obtain them; **iii)** Finally, we obtain a quite simplified solution for this approach, which simplifies to a unidimensional comparison with an optimized threshold. The numerical results indicate a significantly improved performance of our proposed BC approach compared to other protocols available in literature, besides of showing its robustness to the variation of different system parameters.

II. System Model

We consider a set of UEs served by a BS with M antennas, with time and frequency resources divided into coherence blocks of T channel uses. The BS is located in the center of a hexagonal cell and operates through a time-division duplex (TDD), similarly as in [2]. We define $\mathcal{U}_i$ as the set of all UEs inside cell i, and $\mathcal{A}_i$ as the subset of all active UEs, such that $\mathcal{A}_i \subset \mathcal{U}_i$. In a typical overcrowded scenario, we have $|\mathcal{U}_i| \gg T$. However, there is a probability $P_a \leq 1$ that inactive UEs want to become active. Hence, we can consider a scenario where $|\mathcal{A}_i| < T$, in which the BS can temporarily make orthogonal payload data pilot (PDP) signals available to all active UEs during payload data transmission, by employing a grant-based RA protocol.

In this letter, the investigated system model focuses on a central cell arbitrarily chosen and called cell 0, in which, $\mathcal{K}_0 = \mathcal{U}_0 \backslash \mathcal{A}_0$ is the set of inactive UEs with cardinality $K_0 = |\mathcal{K}_0|$. The channel vector between BS and UE k is denoted by $\mathrm{h}_k \in \mathbb{C}^{M\times 1}$. The channel follows a complex Gaussian distribution $\mathrm{h}_k \sim \mathcal{CN}(0, \beta_k \mathrm{I}_M)$, where β_k is the large-scale fading coefficient, obtained as in [2]. The K_0 UEs share τ_p orthogonal RA pilot signals: $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{\tau_p}$, with $\boldsymbol{\psi}_t \in \mathbb{C}^{\tau_p}$ and $||\boldsymbol{\psi}_t||^2 = \tau_p$, $t \in \{1, 2, \ldots, \tau_p\}$. UEs that want to become active randomly choose one of the τ_p RA pilot signals and make an access attempt by transmitting $\boldsymbol{\psi}_{c(k)}$ with power $\rho_k > 0$, with $c(k) \in \{1, 2, \ldots, \tau_p\}$. The set $\mathcal{S}_t = \{k : c(k) = t, \rho_k > 0\}$ contains the indices of the UEs that transmit the pilot t, and thus $|\mathcal{S}_t|$ represents the number of UEs that choose the pilot $\boldsymbol{\psi}_t$ and follows a binomial distribution [2]:

$$|\mathcal{S}_t| \sim \mathcal{B}\left(K_0, \frac{P_a}{\tau_p}\right). \tag{1}$$

The SUCRe protocol is divided into four steps. In the first step, the BS receives the signal $\mathrm{Y} \in \mathbb{C}^{M\times\tau_p}$ from the pilots sent by the UEs:

$$\mathrm{Y} = \sum_{k\in\mathcal{K}_0} \sqrt{\rho_k}\, \mathrm{h}_k\, \psi_{c(k)}^T + \mathrm{W} + \mathrm{N}, \tag{2}$$

in which $\mathrm{N} \in \mathbb{C}^{M\times\tau_p}$ is the noise matrix of the signal received by the BS with each element following a complex normal distribution, $\mathcal{CN}(0, \sigma^2)$, and $\mathrm{W} \in \mathbb{C}^{M\times\tau_p}$ is the interference from adjacent cells. Then BS correlates Y with $\boldsymbol{\psi}_t$ and obtains

$$\begin{aligned} \mathrm{y}_t = \mathrm{Y}\,\frac{\boldsymbol{\psi}_t^*}{||\boldsymbol{\psi}_t||} &= \sum_{i\in\mathcal{S}_t} \sqrt{\rho_i}||\boldsymbol{\psi}_t||\,\mathrm{h}_i + \mathrm{W}\,\frac{\boldsymbol{\psi}_t^*}{||\boldsymbol{\psi}_t||} + \mathrm{n}_t \\ &= \sum_{i\in\mathcal{S}_t} \sqrt{\rho_i\tau_p}\,\mathrm{h}_i + \mathrm{W}\,\frac{\boldsymbol{\psi}_t^*}{||\boldsymbol{\psi}_t||} + \mathrm{n}_t, \end{aligned} \tag{3}$$

in which $\mathrm{n}_t = \mathrm{N}\,\frac{\boldsymbol{\psi}_t^*}{||\boldsymbol{\psi}_t||}$ is the effective noise with distribution $\mathcal{CN}(0, \sigma^2\,\mathrm{I}_M)$.

Then, in step 2 the BS responds to all UEs who sent pilot signals with a precoded signal $\mathrm{V} \in \mathbb{C}^{M\times\tau_p}$:

$$\mathrm{V} = \sqrt{q}\sum_{t=1}^{\tau_p} \frac{\mathrm{y}_t^*}{||\,\mathrm{y}_t\,||}\boldsymbol{\psi}_t^T, \tag{4}$$

in which q is the signal power available at this stage by the BS. The k-th UE then receives the signal $\mathrm{z}_k \in \mathbb{C}^{\tau_p}$:

$$\mathrm{z}_k^T = \mathrm{h}_k^T\,\mathrm{V} + \boldsymbol{\nu}_k^T + \boldsymbol{\eta}_k^T, \tag{5}$$

where $\boldsymbol{\nu}_k^T \in \mathbb{C}^{\tau_p}$ is inter-cellular interference (ICI) and $\boldsymbol{\eta}_k^T$ is the noise of the signal received by the UE in this step, with distribution $\mathcal{CN}(0, \sigma^2\,\mathrm{I}_{\tau_p})$. Then the UE correlates z_k with its chosen pilot $\boldsymbol{\psi}_t$, resulting in

$$z_k = \mathrm{z}_k^T\,\frac{\psi_t^*}{||\psi_t||} = \sqrt{q\tau_p}\,\mathrm{h}_k^T\,\frac{\mathrm{y}_t^*}{||\,\mathrm{y}_t\,||} + \nu_k^T\,\frac{\psi_t^*}{||\psi_t||} + \eta_k, \tag{6}$$

where $\eta_k \sim \mathcal{CN}(0, \sigma^2)$. Defining α_t as:

$$\alpha_t = \sum_{i\in\mathcal{S}_t} \rho_i\beta_i\tau_p + \omega_t, \tag{7}$$

which is the sum of the signal strengths and interference received by the BS during the first stage of the protocol for each pilot t, according to eq. (3). As proposed in [2], the value of α_t can be estimated by $\hat{\alpha}_{t,k}$:

$$\hat{\alpha}_{t,k} = \max\left(\left[\frac{\Gamma(M+\frac{1}{2})}{\Gamma(M)}\right]^2 \frac{q\rho_k\beta_k^2\tau_p^2}{[\Re(z_k)]^2} - \sigma^2, \rho_k\beta_k\tau_p\right), \tag{8}$$

in which $\Re(\cdot)$ returns the real part of a complex number and $\Gamma(\cdot)$ is the gamma function.

In the third step, each UE k knowing his own average channel gain β_k decides whether to retransmit the pilot signal or not. The objective here is to let only one contending UE (the strongest one) retransmit to resolve the collision and connect this UE to the network. The pilot will be retransmitted when $\mathcal{R}_k$ is true and will not be retransmitted when $\mathcal{I}_k$ is true:

$$\mathcal{R}_k : \rho_k\beta_k\tau_p > \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{9}$$

$$\mathcal{I}_k : \rho_k\beta_k\tau_p \leq \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{10}$$

where ϵ_k is a bias parameter with a suitable value proposed in [2]. The fourth step of the SUCRe method consists of granting access to network resources to the UE that successfully retransmits its pilot when the collision has been resolved.

The soft-SUCRe (s-SUCRe) protocol of [5] follows a similar framework. However, instead of applying the hard decision retransmission criterion described in (9) and (10), each UE retransmits or not its chosen pilot in the third step according to its probability of being the strongest contender, which is derived in [5, eqs. (16) and (17)].

III. Bayesian Classifier

We present in this section how the Bayesian classifier can be applied to the random access problem in crowded massive MIMO networks, under the strongest user criterion to resolve pilot collisions. First, in order to obtain an input variable for the proposed BC, we rearrange inequations (9) and (10) as

$$\mathcal{R}_k : \phi_k > 0.5, \qquad \mathcal{I}_k : \phi_k \leq 0.5, \tag{11}$$

in which the $\phi_k \in [0, 1]$ factor represents the proportion of the signal strength of the k-th UE among the contenders:

$$\phi_k = \frac{\rho_k \beta_k \tau_p - \epsilon_k}{\hat{\alpha}_{t,k}}. \tag{12}$$

The problem of resolving the pilot collisions following the strongest UE retransmission rule can be seen as a classification problem to be solved at the UEs' side. We can thus define two classes for the UEs: C_0 represents the class of UEs which are not the strongest contenders for their chosen pilots, and C_1 represents the class of the strongest UEs, forming the set $\mathcal{C} = \{C_0, C_1\}$. We also define the state of each UE k as $\Omega_k \in \mathcal{C}$. Given the ϕ_k estimate obtained after step 2 as in (12), the proposed BC operates by seeking the class that maximizes the *a posteriori* probability of the k-th UE as

$$\begin{aligned} \widehat{\Omega}_k &= \underset{C_\ell \in \mathcal{C}}{arg\,max} \quad p(C_\ell|\phi_k), \\ &= \underset{C_\ell \in \mathcal{C}}{arg\,max} \quad \frac{p(\phi_k|C_\ell)P(C_\ell)}{p(\phi_k)}, \\ &= \underset{C_\ell \in \mathcal{C}}{arg\,max} \quad p(\phi_k|C_\ell)P(C_\ell). \end{aligned} \tag{13}$$

Based on the possibility of the existence of UEs $\in C_1$ with $\phi_k \leq 0.5$, the Bayesian method is then proposed to map and classify the values of ϕ_k statistically. However, as the probabilities in (13) are difficult to be obtained analytically, since the estimate ϕ_k is computed as the ratio of two random variables in (12), we conduct in the following empirical statistical analysis, aiming to numerically obtain the joint PDFs $p(\phi_k|C_\ell)P(C_\ell)$, $\ell = 0, 1$.

The training data of the BC are generated in the software MATLAB 2020a in a simulation setup following that proposed and publicly shared by the authors of the SUCRe protocol [2]. The setup considers the system parameters indicated in the beginning of Section IV. The offline training procedure we conducted consists in generating 10×10^6 examples of ϕ_k values labeled in terms of C_ℓ, $\ell = 0, 1$, equally distributed between the different K_0 values. Then, we apply the MATLAB function *histcounts* configured with 38-bin divisions and `probability` normalization[1] on the training data. The procedure is initially performed without ICI, obtaining Fig. 1, and then repeated with ICI, obtaining Fig. 2.

Fig. 1 presents the estimate of the PDFs of the ϕ_k values conditioned to the occurrences of each class, weighted by the class probabilities, i.e., $p(\phi_k|C_\ell)P(C_\ell)$, $\ell = 0, 1$, for the scenario without ICI. The Figure also depicts the crossover point between both curves, pointing out the ϕ_k value at which $p(\phi_k|C_1)P(C_1)$ becomes greater than $p(\phi_k|C_0)P(C_0)$, which occurs for $\phi_k \approx 0.2942$. Similarly,

[1]This implies that the area under graphs (a) and (b) of the figures add up to one, since $\int_0^1 p(\phi_k, C_0)d\phi_k + \int_0^1 p(\phi_k, C_1)d\phi_k = 1$. Besides, the choice for a number of examples of 10×10^6 was empirical, trading-off a training execution complexity of few minutes while obtaining smooth PDF curves.

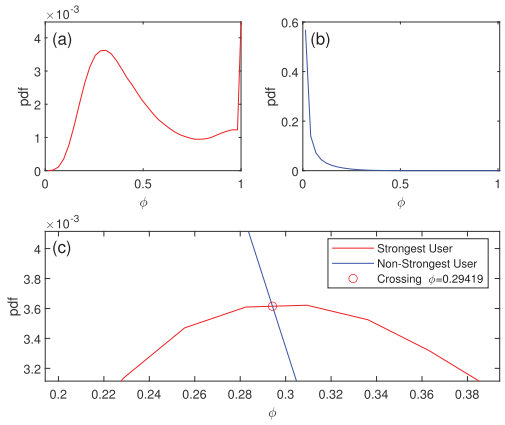


Fig. 1. $p(\phi_k|C_\ell)P(C_\ell)$, (a) $\ell = 1$, (b) $\ell = 0$, and (c) the crossover point for the scenario without ICI.

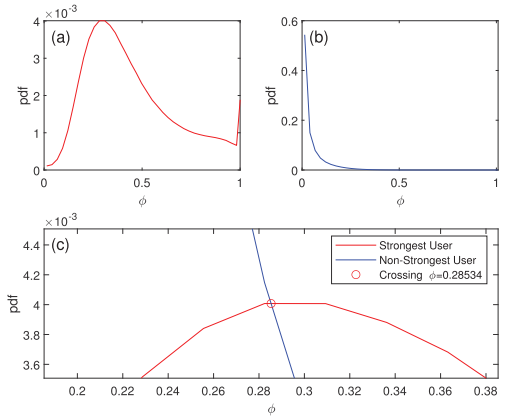


Fig. 2. $p(\phi_k|C_\ell)P(C_\ell)$, (a) $\ell = 1$, (b) $\ell = 0$, and (c) the crossover point for the scenario with ICI.

Fig. 2 presents a similar result for ICI cases. In this scenario, the crossover point turns to occur for $\phi_k \approx 0.2853$.

Interestingly, one can see from the above analysis that a BC applied as criterion to the UEs retransmit their pilots in step 3 of the RA protocol is equivalent to changing the retransmission rule decision threshold of 0.5 from the SUCRe protocol to a more refined value of 0.2942 in the scenario without ICI, or 0.2853 in the scenario with ICI. Therefore, inequations in (11) can be replaced by

$$\begin{aligned} & \text{(w/o ICI)} & \text{(with ICI)} & \\ \mathcal{R}_k : \ & \phi_k > 0.2942, & \phi_k > 0.2853, & \quad (14) \\ \mathcal{I}_k : \ & \phi_k \leq 0.2942, & \phi_k \leq 0.2853. & \quad (15) \end{aligned}$$

In other words, although the purpose of SUCRe protocol is to resolve pilot collisions letting only the strongest UE retransmit their pilots, it takes this decision evaluating $\phi_k > 0.5$. Indeed, $\phi_k > 0.5$ is a sufficient condition for a UE being the strongest contender, but not necessary, giving rise to excessive occurrences of false-negatives, deteriorating connectivity performance. On the other hand, our proposed BC applies a more refined decision threshold near 0.29, which significantly improves the RA performance as shown in the following section.

TABLE I
CONFUSION MATRIX OF SUCRE W/O ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	**6829732** 93.1%	**293539** 4.0%	95.9% 4.1%
C_1	**2977** 0.0%	**210311** 2.9%	98.6% 1.4%
total	100.0% 0.0%	41.7% 58.3%	96.0% 4.0%

TABLE II
CONFUSION MATRIX OF PROPOSED BC W/O ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	**6763743** 92.2%	**127111** 1.7%	98.2% 1.8%
C_1	**68966** 0.9%	**376739** 5.1%	84.5% 15.5%
total	99.0% 1.0%	74.8% 25.2%	97.3% 2.7%

IV. NUMERICAL RESULTS

The performance of the investigated RA protocols are numerically evaluated in this section. The setup considers a hexagonal cell with a radius of 250 m and the number of UEs K_0 varying between 100 and 30000 in increments of 100. The cell is surrounded by six neighboring hexagonal cells with the same radius, but with a fixed number of 10 active UEs. The system operates under the 5G sub-6 GHz band, and the simulation parameters are set as M=100, $\rho = 1, q = 1$, $\tau_p = 10$, and transmission probability $P_a = 0.001$. First, we provide results for the classification accuracy of the schemes, and then we present and compare the RA performance of the protocols.

A. Classification Performance

Tables I and II evaluate the classification performance for the scenario without ICI in terms of confusion matrices. Table I shows the results when considering a decision threshold of $\phi_k = 0.5$, which is employed in the SUCRe protocol, while Table II shows the results when employing the proposed BC approach, with the obtained new threshold of $\phi_k = 0.2942$. The third column at the far right of the matrix shows the precision of predictions for each state belonging to $\widehat{\Omega}_k$. In other words, it presents the accuracy of the classifier given each output class $\widehat{\Omega}_k = C_0$ or $\widehat{\Omega}_k = C_1$. The bottom row shows the successful classification rates of each state. The far-right square at the bottom shows the overall accuracy[2] of the classifier. In Table I the results indicate that for the decision threshold of $\phi_k = 0.5$ employed in the SUCRe protocol, the classification success rate of UEs of class C_0 is 100%. However, the successful classification rate among C_1 UEs is 41.7%, indicating too many false-negatives when $\phi_k \leq 0.5$. Besides, the precision of C_0 and C_1 classifications are 95.9% and 98.6% respectively. The overall accuracy of successful classifications is 96%.

Table II indicates the results when the decision threshold is $\phi_k = 0.2942$. The classification precision of C_0 and C_1 outputs are 98.2% and 84.5%, respectively. The percentages of correct predictions are 99% for class C_0 UEs and 74.8% for class C_1 UEs. The overall accuracy is 97.3%, higher than the results obtained for the threshold $\phi_k = 0.5$.

[2]The total accuracy is given by $P(\widehat{\Omega}_k = C_0, C_0) + P(\widehat{\Omega}_k = C_1, C_1)$. Each of these probabilities can be computed by dividing the respective number of occurrences with the total number of occurrences simulated.

TABLE III
CONFUSION MATRIX OF SUCRE WITH ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	**6866246** 93.0%	**336464** 4.6%	95.3% 4.7%
C_1	**1844** 0.0%	**179662** 2.4%	99.0% 1.0%
total	100.0% 0.0%	34.8% 65.2%	95.4% 4.6%

TABLE IV
CONFUSION MATRIX OF PROPOSED BC WITH ICI

Predicted Class \ Actual Class	C_0	C_1	total
C_0	**6795980** 92.0%	**140362** 1.9%	98.0% 2.0%
C_1	**72110** 1.0%	**375764** 5.1%	83.9% 16.1%
total	99.0% 1.0%	72.8% 27.2%	97.1% 2.9%

Tables III and IV are the confusion matrices for the cases with ICI. Table III shows the results adopting the decision threshold of $\phi_k = 0.5$. The results indicate accuracy of 95.3% for output class C_0 and 99% for output class C_1. The rate of successful classified states are 100% and 34.8% for UEs belonging to classes C_0 and C_1, respectively. The overall accuracy is 95.4%.The inferior classifier performance is noteworthy when taking the ICI into account.

Table IV shows the results when taking $\phi_k = 0.2853$ as the decision threshold. The results indicate an accuracy of 98.0% for output class C_0 and 83,9% for output class C_1. The rate of successful classified states are 99.0% and 72.8% for UEs belonging to classes C_0 and C_1, respectively. The overall accuracy is 97.1%, slightly inferior than the case without ICI.

In summary, the results in this subsection show that the BC approach proposed in this letter trades-off a very slight

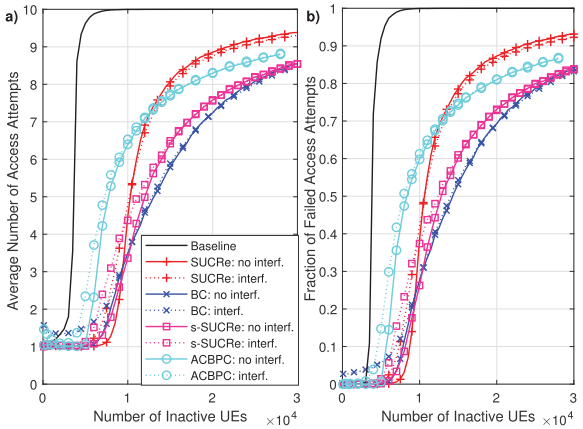


Fig. 3. Performance with K_0 variation: **a)** ANAA, and **b)** FFAA.

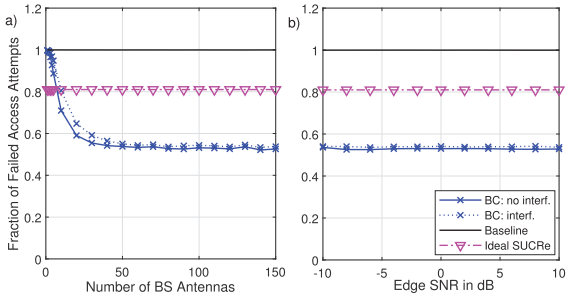


Fig. 4. FFAA with $K_0 = 15000$ UEs *vs* **a)** Number of BS Antennas M, and **b)** Edge SNR. Ideal SUCRe refers to SUCRe with perfect $\hat{\alpha}_{t,k}$ estimation, i.e., $\hat{\alpha}_{t,k} = \sum_{i \in \mathcal{S}_t} \rho_i \beta_i \tau_p$.

increase in the false-positive rates for a significant reduction in the false-negative probabilities in comparison with SUCRe protocol. Consequently, the correct classification for UEs of class C_1 is significantly increased using the BC. The results point to 41.7% of SUCRe to 74.8% in the scenario without ICI and 34.8% to 72.8% in ICI. These results are also very beneficial to the RA performance, as shown in the following.

B. Connectivity Performance

Fig. 3 shows the result of Average Number of Access Attempts (ANAA) and Fraction of Failed Access Attempts (FFAA). The black line represents the performance of the baseline protocol of [2], which is an ALOHA-like protocol where pilot collisions are only handled by retransmission in later RA blocks. The red lines with "+" marker indicate the results obtained with the original SUCRe protocol of [2], the magenta lines with "□" marker indicate the results obtained with the s-SUCRe protocol of [5], and the ciano lines with "○" marker indicate the results obtained with the ACBPC protocol of [6]. The blue lines with "x" marker indicate the results obtained when employing the BC methodology proposed in this letter. The dotted line shows the results obtained for the cases with ICI. Both charts show a significantly high performance of the BC for $K_0 > \tau_p/P_a = 10000$ UEs, when the average number of accessing UEs surpasses the number of available pilots. Compared with s-SUCRe for example, the FFAA decreases 6.73% for $K_0 = 15000$ UEs, while the proposed BC approach avoids the complexity of computing the probability of the UE being the strongest contender, and does not require the actual K_0 value to be known at the UEs' side as in s-SUCRe. Besides, Fig. 4 shows the FFAA with the variation in the number of BS antennas M and edge SNR, defined as ρ/σ^2 and q/σ^2, with $\rho = q$. Although the proposed BC is trained with $M = 100$ and 0 dB of edge SNR, the results show the robustness of the approach regarding the variation of such parameters, since the BC remains achieving improved performances. Fig. 4a also shows that 50 BS antennas are sufficient for a satisfactory operation of the BC approach.

V. Final Remarks

This letter proposes a statistical approach to optimize the SUCRe protocol through a Bayesian classifier. The classification results indicated that our proposed methodology trades-off a slight increase in false-positive rates for a significant reduction in the false-negative probabilities, substantially increasing the correct classification probabilities for the strongest UEs. This behavior shows to be very beneficial from the connectivity performance perspective, which has been significantly improved by our proposed method in comparison with other protocols, for both scenarios with and without ICI. The overcrowded scenario, with more expected accessing UEs than available RA pilots, i.e., $K_0 > \tau_p/P_a$, is where the most significant performance improvements have been found, indicating the proposed method as a promising approach in massive MIMO systems with very high density of UEs.

References

[1] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.

[2] E. Björnson, E. de Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A random access protocol for pilot allocation in crowded massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2220–2234, Apr. 2017.

[3] H. Han, X. Guo, and Y. Li, "A high throughput pilot allocation for M2M communication in crowded massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9572–9576, Oct. 2017.

[4] H. Han, Y. Li, and X. Guo, "A graph-based random access protocol for crowded massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7348–7361, Nov. 2017.

[5] J. C. Marinello and T. Abrão, "Collision resolution protocol via soft decision retransmission criterion," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4094–4097, Apr. 2019.

[6] J. C. Marinello, T. Abrão, R. D. Souza, E. de Carvalho, and P. Popovski, "Achieving fair random access performance in massive MIMO crowded machine-type networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 503–507, Apr. 2020.

[7] H. Han, Y. Li, and X. Guo, "User identity-aided pilot access scheme for massive MIMO-IDMA system," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6197–6201, Jun. 2019.

[8] H. Han, L. Fang, W. Lu, K. Chi, W. Zhai, and J. Zhao, "A novel grant-based pilot access scheme for crowded massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11111–11115, Oct. 2021.

[9] O. S. Nishimura, J. C. Marinello, and T. Abrão, "A grant-based random access protocol in extra-large massive MIMO system," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2478–2482, Nov. 2020.

[10] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.

[11] J.-B. Seo, B. C. Jung, and H. Jin, "Modeling and Online adaptation of ALOHA for low-power wide-area networks (LPWANs)," *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15608–15619, Oct. 2021.

[12] Y. Tao, B. Li, and C. Zhao, "Channel Detection under impulsive noise and fading environments for smart grid," *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15407–15421, Oct. 2021.

APPENDIX B – REINFORCEMENT LEARNING-BASED GRANT-FREE RANDOM ACCESS FOR MMTC MASSIVE MIMO NETWORKS

Reinforcement Learning-Based Grant-Free Random Access for mMTC Massive MIMO Networks

Felipe A. Dutra Bueno, Alessandro Goedtel, Taufik Abrão, José Carlos Marinello Filho

***Abstract*—The expected huge number of connected devices in Internet of Things (IoT) applications characterizes the massive machine-type communication (mMTC) scenario, one prominent use case of beyond fifth-generation (B5G) systems. To meet the mMTC connectivity requirements, *grant-free* (GF) random access (RA) protocols are seen as a promising solution, due to the small amount of data that the simple devices usually transmit. In this letter, we propose a GF RA protocol based on a multi-agent reinforcement learning approach, applied to aid the IoT devices in selecting the least congested RA pilots. The rewards obtained by the devices in collision cases resemble the congestion level of the chosen pilot. To enable the operation of the proposed method in a realistic B5G network scenario, and aiming to reduce signaling overheads and centralized processing, the rewards in our proposed method are computed by the devices taking advantage of the large number of base station antennas. Numerical results demonstrate the superior performance of the proposed method in terms of latency, network throughput, and per-user throughput compared with other protocols.**

***Index Terms*—Random access protocol, Grant-Free, Beyond 5G, Reinforcement Learning, Massive MIMO.**

I. Introduction

Cellular Internet of Things (CIoT) is an important research topic within beyond fifth-generation (B5G) networks [1]. The number of IoT devices has been explosively increasing recently, while most devices are low-power nodes whose batteries are expected to be usable for years. Furthermore, such IoT devices are usually distributed over a long-range. Therefore, we can point out the main requirements of CIoT in B5G networks: massive connectivity, low power consumption, and broad coverage [1]. Exploiting new wireless technologies, such as massive multiple-input multiple-output (MIMO) technology, intelligent reflecting surfaces, and others, is essential to achieve such goals.

The massive MIMO is already a successful technology [2]. Its fundamental idea is to equip base stations (BSs) with many antennas to serve a set of single-antenna users scattered in the cell. It benefits from the fact that the effects of uncorrelated noise and fast fading disappear as the number of BS antennas grows to infinity, remaining only the inter-cellular interference that results from pilot contamination [3]. Massive MIMO systems usually employ a time-division duplex (TDD) scheme that demands the transmission of only uplink (UL) pilot signals to acquire channel state information (CSI), since the downlink (DL) CSI can be estimated by channel reciprocity [4]. However, while the number of devices is continually increasing, motivated by the wide spreading of IoT applications, the number of resources offered by the cellular network remains scarce. This fact gives rise to performance issues such as pilot collisions when two or more users choose the same pilot trying to access the BS resources. Therefore, establishing an effective random-access (RA) policy is mandatory.

Several methods have been presented to enhance the traditional random access performance, such as access class barring (ACB), slotted access, and backoff [5]. Among several proposed solutions for congestion problem resulted from massive access, there are for instance, **a**) [5], which investigates an efficient random access procedure based on ACB to decrease the access delay and the power consumption under wireless networks congestion resulted from massive access; **b**) the *strongest-user collision resolution* (SUCRe) protocol from [6]. The SUCRe protocol is a grant-based (GB), 4-steps RA approach whose main idea consists in allowing just the strongest pilot competitor to access the BS resources each time. Numerical results indicate that the SUCRe protocol can solve about 90% of all collisions. However, the 4-steps procedure required for the BS to grant exclusive communication resources to the devices could result in a performance bottleneck, being a source of excessive delay, and signaling overhead. Therefore, it is not the ideal choice for massive machine-type communication (mMTC) systems, where accessing devices usually have small data packets to transmit sporadically. Other works such as [7]–[10] present proposals for optimization of the SUCRe protocol showing promising results. However, they are also GB protocols that introduce extra complexity or overhead to the SUCRe protocol. B5G RA schemes should achieve high scalability under latency and reliability constraints to support new use cases. For this purpose, grant-free (GF) RA protocols have gained increasing interest, as they can drastically reduce control signaling for connection establishment [1].

Many GF RA protocols available in the literature are derived from the contention resolution diversity slotted ALOHA [11], and irregular repetition slotted ALOHA [12]. The idea behind such schemes is to repeat the transmission of data packets in several randomly-chosen slots, including the indices of the chosen slots as side information. Whenever a particular device chooses a slot, its payload is successfully decoded, and its interference in the other slots is canceled through successive interference cancellation (SIC), increasing the occurrence of other non-colliding slots. Although achieving exceptional performances, these protocols have the *drawbacks* of requiring packets re-transmissions, a substantial overhead for side information signaling, increased complexity for SIC

F. A. D Bueno, A. Goedtel, J. C. Marinello are with Electrical Engineering Department, Federal University of Technology PR, Cornélio Procópio, PR, Brazil (e-mail: faugustobueno@gmail.com, agoedtel@utfpr.edu.br, jcmarinello@utfpr.edu.br).

T. Abrão is with Department of Electrical Engineering Londrina State University, Parana Brazil (e-mail taufik@uel.br)

evaluation, and the possibility of propagation errors.

Reinforcement learning (RL) techniques can be aggregated to deal with *congestion control* (CC), especially in grant-free random access M-MIMO networks. The CC is a fundamental mechanism for the implementation of heterogeneous networks. In the recent years, RL-based smart CC methods have drawn a lot of attention in the research community. RL is a type of machine learning technique, which can enable the agents/devices to interact with the environment in order to learn efficient strategies that maximize the long-term system performance. The most typical RL algorithm is the Q-learning (QL) algorithm, which can be implemented at the user equipment (UE) even without an operating model of the environment.

In ultra-dense mMTC networks, RA optimization is very challenging; hence, the use of ML tools is quite promising, particularly RL, to efficiently accommodate the MTC devices in RA slots. However, exploitation *vs* exploration is critical in reinforcement learning. RL algorithms try to continuously optimize the environment learning for the best returns, attempting to efficiently achieve good exploitation, while the exploration mechanism plays a secondary role. Consequently, RL agents try to find the best solution as fast as possible. However, committing to solutions too quickly without enough exploration could become inadequate, as it could lead to local minima or total failure.

In [13], an RL-based GF RA for pilot collision control is proposed. The RL algorithm used in [13] is the Q-Learning, where each accessing device is an independent agent and the rewards are simply $+1$ when it chooses a unique RA slot or -1 otherwise. A similar procedure is proposed in [14], in which the BS also sends $+1$ or -1 depending on the outcome of the device's transmission. However, in the case of collision, the actual reward computed by the device is -1 times the ratio of packets already transmitted by it, *e.g*, if the device has already transmitted 20% of its packets and collided in the current instant, its actual reward is -0.2. As an improvement, a collaborative QL RA scheme is proposed in [15], in which the negative rewards in case of collisions are proportional to the congestion level of the chosen RA slot. However, the protocol assumes that the devices know the exact number of devices colliding by their chosen RA slot. The performance results are better than when compared with the independent QL approach of [13]. Nonetheless, [13], [14], and [15] do not assume a realistic system model and do not consider channel effects like multipath fading, path loss, thermal noise, and inter-cell interference (ICI), besides assuming the devices know the exact congestion levels in the case of [15].

In this work, we propose a QL-based GF RA protocol specially designed for realistic mMTC B5G scenarios. Our scenario considers realistic wireless propagation effects, including multipath fading, shadowing, path loss, thermal noise, and ICI. Besides, the devices compute the collaborative QL rewards in collision cases by estimating the congestion levels.

Contributions. The paper's contributions are threefold:

i. ***Innovative Grant-free RA protocol***. To suitably operate under such a realistic scenario, and allowing *improved congestion level estimation* at the devices with *minimal signaling overhead*, we propose an *innovative GF RA protocol* taking advantage of the large number of BS antennas typically available in the advanced B5G networks.

ii. ***Improved latency, network throughput, and per-user throughput figures of merit***: numerical results reveal that latency, network throughput, and per-user throughput performances obtained through the congestion estimation are near the ideal case where the device knows the actual congestion level.

iii. ***Robustness against system parameter changes***: furthermore, the network throughput and latency results also show robustness when varying the number of antennas or the number of packets that each device has to send, in a scenario with an equal number of contending devices and RA pilots.

The remainder of this paper is organized as follows. Section II revisited the main QL-based GF RA protocols available in the literature. The proposed 2-step QL GF RA M-MIMO Protocol is diligently described in Section III. Numerical results exploring the main metrics for analysing the performance of random access networks are carried out in Section IV. The main conclusions and possible research directions are offered in Section V.

II. Existing QL-based GF RA Protocols

Q-learning can be used as a multi-agent RL method to aid IoT devices in selecting the least congested RA pilots in a decentralized way. In typical mMTC networks, active IoT devices can coordinate neither with the BS nor with each other for pilot selection; therefore, IoT devices act as individual learning agents, using their previous experience to enhance the probability of selecting exclusive RA pilots, minimizing the occurrence of pilot collisions. In this section, we revisited two previous works where this framework has been applied in a *simple collision channel*, *i.e.*, assuming that the only communication impairment is *pilot collision*. The adopted scenario considers an mMTC network employing a GF RA policy in a cell comprising K_a IoT devices, each with L packets to transmit, and contending for τ_p pilot resources.

The devices transmit the payload data prepended with an RA pilot to enable channel estimation employing a GF RA protocol. If the chosen RA pilot is exclusively chosen by that device (no interference), it is assumed that the payload data packet transmitted by it is successfully decoded at the BS, and the device proceeds to transmit the next data packet. Otherwise, if a collision occurs, the device repeats the transmission of the same packet in the next frame, which increases *latency* and decreases the *per-user throughput* and the *network throughput*. Thus, *reinforcement learning* techniques can be employed to guide the pilots' choice of devices towards the least congested ones, improving the connectivity performance of the network as a whole.

The interaction between the IoT device and the environment can be modeled as a Markov Decision Process (MDP), where at each time step, a device can change its current state $x_t \in X$ to $x_{t+1} \in X$ by taking action $a_t \in A$

based on a transition probability function $f(x_t, a_t, x_{t+1})$ [15]. Depending on its state-action pair, the device is rewarded with $r_{t+1} \in R$ during the transition state. Besides, the expected return of a state-action pair is given by $Q^{\pi}(x,a) = \mathbb{E}\left[\sum_{j=0}^{J} \gamma^j r_{t+j+1} | x_t = x, a_t = a, \pi\right]$, where π is the established policy, $\gamma \in [0,1]$ is the discount factor and J is the length of one episode [15].

Satisfying the Bellman optimality equation, the Q-function can then be written as $Q^*(x,a) = \max_{\pi} Q^{\pi}(x,a)$. If a *greedy policy* $\pi(x) = \arg\max_a Q^{\pi}(x,a)$ is established for the Q-function, then we have a QL algorithm that selects only the actions associated with the highest Q-value $Q(x,a)$ at each state, calculated iteratively as:

$$Q_{t+1}(x_t, a_t) = Q_t(x_t, a_t) + \delta_t \left[r_{t+1} + \gamma \max_a Q_t(x_{t+1}, a) - Q_t(x_t, a_t) \right], \quad (1)$$

where δ_t is the learning rate at the tth time step. This model can then be applied for *decentralized pilot selection* as described in [13], [15], in which each device has a Q-table of τ_p elements evaluating its experience in selecting the different pilots to transmit data. This Q-table can be updated according to:

$$Q_{t+1}(k,\ell) = Q_t(k,\ell) + \delta \left[R(k,\ell) - Q_t(k,\ell) \right], \quad (2)$$

where $Q_t(k,\ell)$ is related to the experience of the kth device in choosing the pilot $\ell = c(k)$, and $R(k,\ell)$ indicates the reward function for this choice.

A simple way to compute the rewards $R(k,\ell)$ in (2) is proposed in [13] through an *independent* QL approach for mMTC (**iQmMTC**), where $R(k,\ell) = +1$ if the transmission succeeds, or $R(k,\ell) = -1$, otherwise. However, [15] has shown that better results can be achieved through a *collaborative* approach (**cQmMTC**), where the device is rewarded either with $R(k,\ell) = +1$ for a successful transmission or with $R(k,\ell) = -P_c(k)$ if the transmission fails, with the collaborative penalty function $P_c(k)$ being computed as

$$P_c(k) = \frac{1}{K_a} CL(k), \quad (3)$$

where K_a is the number of active devices in the cell and the congestion level $CL(k) = |\mathcal{S}_{c(k)}|$ is the *number of contenders* for pilot $c(k)$ [15], while $\mathcal{S}_{c(k)}$ is the set of devices choosing the same pilot $c(k)$.

In this framework, each device keeps its own $1 \times \tau_p$ Q-table. Initially, this Q-table is filled with zeros, and all τ_p RA pilots are equally like to be chosen by it. Then, at each subsequent transmission attempt, the devices are rewarded with $R(k,\ell)$, and their Q-tables are updated according to (2). Once an individual device has updated its Q-table, it will only choose pilots among the ones with the highest Q-value. The process is repeated until the L packets are transmitted.

In the cQmMTC approach of [15], the congestion levels in case of collisions are computed as the number of contending devices, $|\mathcal{S}_{c(k)}|$. However, nothing is discussed about the feasibility of making this information available on the device's side. At first glance, it would require to conceive an estimator to be employed at the BS, and then feedback the result to the devices, which would spend significant signaling overhead. For comparison purpose, herein, we assume perfectly known the congestion level at device side, treating this scheme as *genie* cQmMTC. Furthermore, in a practical network, not all IoT devices are active at a given instant, since they activate independently with certain probabilities, in such a way that the actual number of active devices is not known by the BS. To avoid complex computations and excessive signaling overheads while simplifying the procedure, we propose a 2-step QL-based GF RA protocol making use of the large number of BS antennas to allow a collaborative penalty function computation at the devices' side in case of collisions with minimal complexity and overhead.

III. Proposed 2-step QL GF RA M-MIMO Protocol

We consider a TDD scheme, where the channels are considered constant during a time slot. The BS is equipped with a massive number of BS antennas (M), localized at the center of the cell. Let $\mathcal{K}$ be the set of single-antenna devices in the cell, which decide to activate with probability P_a transmitting *payload data* together with a randomly selected *pilot sequence* to enable UL channel estimation at the BS side. An illustrative representation of the adopted scenario is presented in Fig. 1. We consider τ_p mutually orthogonal pilot sequences: $\boldsymbol{s}_1, ..., \boldsymbol{s}_{\tau_p} \in \mathbb{C}^{\tau_p \times 1}$, such that each pilot has length τ_p and $\|\boldsymbol{s}_t\|^2 = \tau_p$, $\forall t \in [1, \tau_p]$. Each device that activates has a number of L_k packets to transmit. Therefore, considering $\mathcal{S}_t \subset \mathcal{K}$ as the set of devices that want to transmit data selecting pilot t, its cardinality follows a binomial distribution:

$$|\mathcal{S}_t| \sim \mathcal{B}\left(K, \frac{P_a}{\tau_p}\right) \quad (4)$$

where $K = |\mathcal{K}|$ is the total number of devices in the cell.

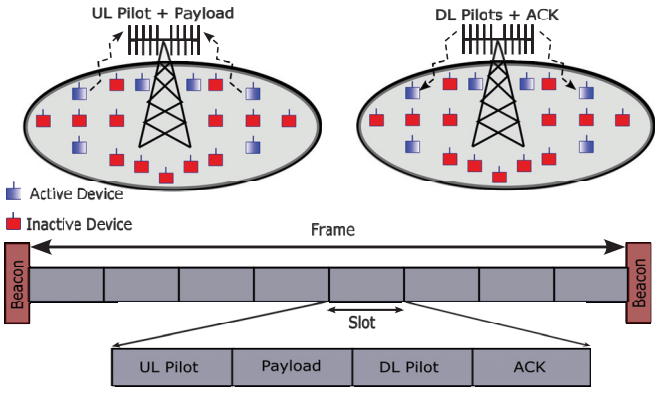


Fig. 1. Illustrative representation of the adopted scenario.

The channel vector between BS and device k is denoted by $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$. The channel follows a complex Gaussian distribution $\mathbf{h}_k \sim \mathcal{CN}(0, \beta_k \mathbf{I}_M)$, where β_k is the large-scale fading coefficient, which follows an urban micro scenario [16]. So the large-scale fading of the link between device k and the BS is

$$\beta_k = 10^{-\kappa \log(d_k) + \frac{g+\varphi}{10}}. \quad (5)$$

In this equation d_k is the distance between device k and the BS, $\kappa = 3.8$ is the path loss exponent, $\varphi \sim \mathcal{N}(0, \sigma_{\text{sf}}^2)$ is

the shadow fading, with standard deviation $\sigma_{\text{sf}} = 10$ dB, and $g = -34.53$ dB is the path loss at the reference distance [16].

When the device k transmits data, it randomly selects one of the τ_p pilot sequences and transmits it followed by its UL payload data packet $\mathbf{d}_k \in \mathbb{C}^{\tau_d \times 1}$, with a non-zero transmit power $\rho_k > 0$, where τ_d is the data length. We can denote the chosen pilot as $c(k) \in \{1, 2, \ldots, \tau_p\}$, and define the UL signal as $\mathbf{x}_k = [\boldsymbol{s}_{c(k)}^T, \mathbf{d}_k^T]^T \in \mathbb{C}^{(\tau_p+\tau_d)\times 1}$.

Thus, the BS receives the signal

$$\mathbf{Y} = [\mathbf{Y}_p, \mathbf{Y}_d] = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k \mathbf{x}_k^T + \mathbf{N}, \tag{6}$$

where $\mathbf{Y} \in \mathbb{C}^{M\times(\tau_p+\tau_d)}$, $\mathbf{Y}_p \in \mathbb{C}^{M\times\tau_p}$, $\mathbf{Y}_d \in \mathbb{C}^{M\times\tau_d}$, and $\mathbf{N} \in \mathbb{C}^{M\times(\tau_p+\tau_d)}$ is the receiver noise with entries drawn from $\mathcal{CN}(0, \sigma^2)$. Besides, we have that

$$\mathbf{Y}_p = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k \boldsymbol{s}_{c(k)}^T + \mathbf{N}_p, \text{ and} \tag{7}$$

$$\mathbf{Y}_d = \sum_{k\in\mathcal{K}} \sqrt{\rho_k}\mathbf{h}_k \mathbf{d}_k^T + \mathbf{N}_d, \tag{8}$$

with $\mathbf{N} = [\mathbf{N}_p, \mathbf{N}_d]$. Hence, the BS correlates (7) with each pilot to generate channel estimates. For the case of an arbitrary pilot $\boldsymbol{s}_t$, with $t \in [1, \tau_p]$, it yields:

$$\mathbf{y}_t = \mathbf{Y}_p \frac{\boldsymbol{s}_t^*}{\|\boldsymbol{s}_t\|} = \sum_{i\in\mathcal{S}_t} \sqrt{\rho_i\tau_p}\mathbf{h}_i + \mathbf{n}_t, \tag{9}$$

where $\mathbf{n}_t = \mathbf{N}_p \frac{\boldsymbol{s}_t^*}{\|\boldsymbol{s}_t\|}$ is the effective receiver noise, so that $\mathbf{n}_t \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_M)$. As a result, the BS tries to decode the payloads in (8) using the channel estimates $\mathbf{y}_t$, evaluating:

$$\widehat{\mathbf{d}}_k^T = \frac{\mathbf{y}_t^H}{\sqrt{\tau_p}}\mathbf{Y}_d. \tag{10}$$

The signal-to-interference-plus-noise ratio (SINR) of $\widehat{\mathbf{d}}_k$ in (10) can be obtained following the SINR analysis of [17], adapting the results to our scenario, as follows:

$$\gamma_k^{\text{ul}} = \frac{M\,\rho_k^2\,\beta_k^2}{M\sum_{\substack{i\in\mathcal{S}_t \\ i\neq k}} \rho_i^2\,\beta_i^2 + \left[\sum_{i\in\mathcal{S}_t} \rho_i\beta_i + \frac{\sigma^2}{\tau_p}\right]\left[\sum_{j\in\mathcal{K}} \rho_j\beta_j + \sigma^2\right]}. \tag{11}$$

We assume that the decoding of $\widehat{\mathbf{d}}_k$ in (10) is always successful when k is the unique competitor for the pilot t (without pilot collisions). The BS responds with an ACK feedback message if the decoding of (10) is successful, together with the transmission of a precoded DL pilot signal $\mathbf{V} \in \mathbb{C}^{M\times\tau_p}$, with power q, according to:

$$\mathbf{V} = \sqrt{\frac{q}{\tau_p}} \sum_{t=1}^{\tau_p} \frac{\mathbf{y}_t^*}{||\mathbf{y}_t||} \boldsymbol{s}_t^T. \tag{12}$$

The devices receive $\mathbf{z}_k \in \mathbb{C}^{\tau_p\times 1}$, $k \in \mathcal{S}_t$

$$\mathbf{z}_k^T = \mathbf{h}_k^T\mathbf{V} + \boldsymbol{\eta}_k^T, \tag{13}$$

where $\boldsymbol{\eta}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{\tau_p})$ is the noise. After correlating $\mathbf{z}_k$ with $\boldsymbol{s}_t$, the device calculates

$$z_k = \mathbf{z}_k^T \frac{\boldsymbol{s}_t^*}{||\boldsymbol{s}_t||} = \sqrt{q}\,\mathbf{h}_k^T \frac{\mathbf{y}_t^*}{||\mathbf{y}_t||} + \eta_k, \tag{14}$$

where $\eta_k \sim \mathcal{CN}(0, \sigma^2)$.

Collaborative penalty function computation. Let $\alpha_t = \sum_{i\in\mathcal{S}_t} \rho_i\beta_i\tau_p$ be the *sum of average channel gains of the devices* in $\mathcal{S}_t$ seen at the BS according to (9), then an asymptotically *error-free estimator* for α_t is proposed in a similar scenario[1] in [6] as:

$$\hat{\alpha}_{t,k} = \max\left(\left[\frac{\Gamma(M+\frac{1}{2})}{\Gamma(M)}\right]^2 \frac{q\rho_k\beta_k^2\tau_p}{[\Re(z_k)]^2} - \sigma^2,\ \rho_k\beta_k\tau_p\right), \tag{15}$$

$\Re(\cdot)$ is the real part and $\Gamma(\cdot)$ is the complete Gamma function.

Given the estimate $\hat{\alpha}_{t,k}$ in (15), reminding that $\alpha_{c(k)} = \sum_{i\in\mathcal{S}_{c(k)}} \rho_i\beta_i\tau_p$, and since the device k knows its average channel gain β_k, it can compute a measure of how congested is its chosen pilot as follows

$$\widehat{\phi}_k = \frac{\hat{\alpha}_{t,k}}{\rho_k\beta_k\tau_p}, \qquad 1 \leq \widehat{\phi}_k < \infty. \tag{16}$$

One can note that as long as $\widehat{\phi}_k$ approaches 1, it indicates to be likely that no other device has chosen the pilot $c(k)$. On the other hand, as $\widehat{\phi}_k$ increases, it indicates to be likely that many other devices chose the same pilot $c(k)$. Therefore, $\widehat{\phi}_k$ can be seen as a *rough estimate* of $CL(k)$; hence, the penalty function $P_c(k)$ in (3) can be computed approximately as

$$P_c(k) \approx \frac{\widehat{\phi}_k}{\widehat{K}_a}, \tag{17}$$

in which $\widehat{K}_a$ is an estimate for the number of active devices in the cell. We propose to employ here a simple estimator, which computes $\widehat{K}_a$ as the expected number of active devices supposing that no pilot collision occurs. In this way, one can compute:

$$\widehat{K}_a = K \cdot P_a \cdot \mathbb{E}[L_k]. \tag{18}$$

The penalty function $P_c(k)$ in (17) can be used in a realistic massive MIMO scenario, allowing a practical implementation of the cQmMTC approach in a GF RA protocol. Therefore, in the proposed protocol, the devices choose their RA pilots along the transmissions of the L_k packets according to their own Q-table, which is updated following (2), while computing the rewards as:

$$R(k,\ell) = \begin{cases} +1, & \text{if the transmission succeeds} \\ -P_c(k) = -\widehat{\phi}_k/\widehat{K}_a, & \text{in case of pilot collision.} \end{cases} \tag{19}$$

IV. Numerical Results

In this section, we evaluate the performance of the proposed QL-based GF RA protocol in terms of: (**i**) average latency, considered herein as the total number of attempts, A_k, the device makes to transmit its L_k packets, (**ii**) average network throughput, defined as the ratio between the number of successfully transmitted packets (without collisions) at certain time step and the number of available pilots τ_p, and (**iii**)

[1] In [6], the $\hat{\alpha}_{t,k}$ estimate is computed as part of the 4-steps GB handshake procedure of SUCRe protocol, and used to let the UEs decide whether they should retransmit the chosen pilot or not, depending if it is the strongest contender. Differently, herein we employ a GF RA protocol, in which the devices transmit the payload data together with an RA pilot to enable channel estimation and data decoding in a reduced number of steps.

the per-user throughput, considered as the ratio of the total number of successfully sent packets by each device, L_k, to the total number of attempts, A_k, that the device has to make to send them, in such a way that $L_k \leq A_k$. For the simulations, we consider a massive MIMO BS equipped with M antennas at the center of a hexagonal cell with a radius of 250m, surrounded by six neighboring hexagonal cells with the same radius. Each neighboring cell has a fixed number of $K_{\text{ici}} = 400$ active interfering devices. The simulation parameters are set as $\rho = 1$, $q = 10$, $\tau_p = 400$, and $\delta = 0.1$. With respect to L and P_a, we investigate three different scenarios in this section: (**i**) $L_k = L$, $\forall k \in \mathcal{K}$, and $P_a = 1$, such that $K_a = K$; (**ii**) random L_k, and $P_a = 1$, such that $K_a = K$; (**iii**) random L_k, $\forall k \in \mathcal{K}$, and $P_a = 0.1\%$, such that $K_a \leq K$ is also random.

We investigate 4 protocols: the **a**) *baseline* scheme, which is equivalent to the slotted ALOHA protocol, with the devices choosing the pilots uniformly at random; **b**) *iQmMTC* approach of [13]; **c**) *cQmMTC* approach of [15], assuming that the actual values of $|\mathcal{S}_t|$ and K_a are perfectly known at the devices' side, like if a genie could inform this to them; and **d**) our proposed 2-step QL GF RA protocol leveraging the *massive* MIMO propagation features to efficiently compute the negative rewards of the QL framework at the devices' side, which we denote as *mQmMTC*. Besides, for this last one, we investigate its performance in the scenarios with and without ICI[2].

A. Fixed L_k and $P_a = 1$

In this subsection, we take $L_k = L$, $\forall k \in \mathcal{K}$, and $P_a = 1$, such that $K_a = K$. Although these simplifying assumptions usually do not hold in practice, they are useful to unveil the full potential of the investigated methods, as well as to evaluate the performance losses when not assuming them, which is carried out in the following subsections. The graphs presented in the Fig. 2 and 3 have been generated with 10000 Monte-Carlo realizations. Each realization is a frame, or a time step in the QL framework, in which each device can transmit only one pilot and the payload packet. The number of antennas is kept fixed at $M = 100$, and the number of devices varies from 25 to 800 in steps of 25.

Fig. 2 shows average latency *versus* K results. The proximity of the proposed mQmMTC results for both scenarios with and without interference with the ideal cQmMTC protocol is noteworthy. Also, both the mQmMTC and the ideal cQmMTC results are below the baseline for any value of K. Also, they are below the independent QmMTC for $K > 400$, which corroborates with the cQmMTC superiority presented in [15]. Similarly, Fig. 3 reveals average network throughput versus K results where the mQmMTC is very close to the ideal cQmMTC performance for both the scenarios with and without interference, while being consistently superior to the baseline results for any number of devices K. Furthermore, compared with iQmMTC, the obtained performances are very similar in the region of $K \leq \tau_p = 400$ devices, while the performance obtained by the proposed mQmMTC approach becomes remarkably superior for a higher number of devices. For example, with $K = 600$ active devices, while baseline and iQmMTC achieve network throughputs of ≈ 0.34 and 0.36, respectively, our mQmMTC protocol achieves a network throughput of ≈ 0.49, an improvement of $\approx 44\%$ and $\approx 36\%$, respectively.

[2]It is worth to note that for the baseline, iQmMTC, and cQmMTC in the adopted scenario, only pilot collisions degrade their connectivity performance. Therefore, ICI does not matter for them.

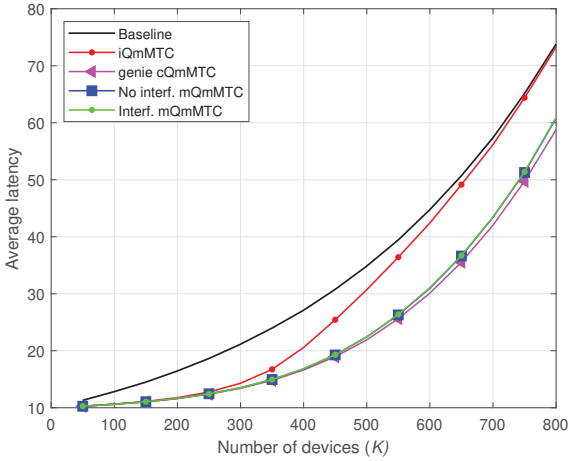


Fig. 2. Average latency $\times K$, for $L = 10$ packets, and $M = 100$.

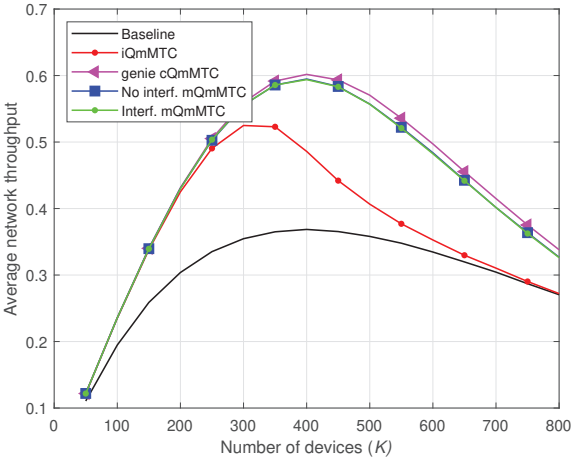


Fig. 3. Average network throughput $\times K$, for $L = 10$ packets, and $M = 100$.

The graphs presented in Fig. 4 and 5 are generated with 80000 Monte-Carlo realizations. The number of devices is fixed at $K = 400$ and the number of available pilots is also fixed at $\tau_p = 400$. The number of BS antennas varies in the range $M \in [1, 100]$, both in the home cell as well as in the neighboring cells. Fig. 4 depicts the average latency with an increasing number of BS antennas M, while Fig. 5 reveals the behavior of the average network throughput *vs.* M. Both figures of merit for the proposed mQmMTC approach improve with the increasing number of antennas M, since the reward computations in (17) benefit from the large number of BS antennas (favorable propagation effect).

The results of the ideal cQmMTC are also included in the figures as a lower bound (avg. latency) and upper bound (avg. throughput), respectively. The average percentual degradations of the results of mQmMTC regarding the ideal cQmMTC is also shown in both figures for $M = 30$ and $M = 100$. One can see that the improvement caused by increasing M from $M = 10$ to $M = 100$ is not as significant as when increasing M from $M = 1$ to $M = 10$. Therefore we can conclude that our proposed mQmMTC RA protocol is able to achieve very improved connectivity performance even with a small number of BS antennas. Indeed, $M \approx 30$ antennas at the BS reveals to be sufficient to attain reliable congestion level estimation $\widehat{\phi}_k$ when a maximum acceptable degradation level of 3.5% is considered.

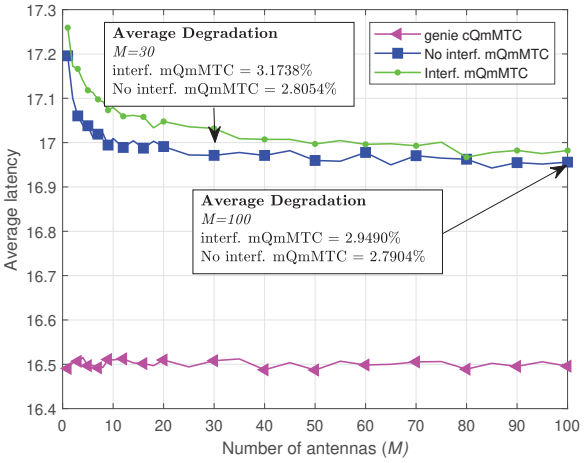


Fig. 4. Average latency $\times M$, for $L = 10$ packets, and $K = 400$.

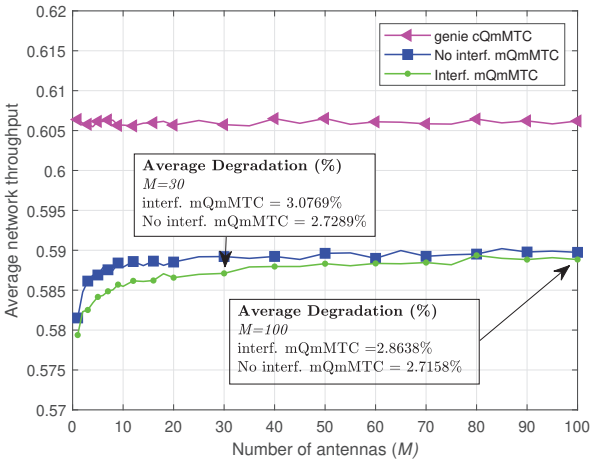


Fig. 5. Average network throughput $\times M$, for $L = 10$ packets, and $K = 400$.

Figures 6 and 7 present, respectively, the graphs of latency and network throughput *versus* L . The results shown in both figures are generated with 10000 Monte-Carlo realizations. The number of active devices are kept fixed at $K = 600$ and the number of antennas is also fixed at $M = 100$. The graphs illustrate the superiority of the proposed methods even when each device has a small number of packets to send ($L \leq 10$). In fact, a minimum number of $L = 2$ is enough for the proposed methods (mQmMTC without and with ICI) to produce a result superior to the baseline and the iQmMTC methods, while approximating to the cQmMTC method.

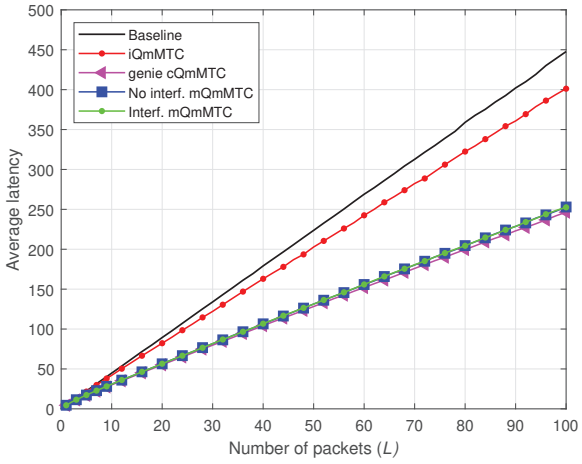


Fig. 6. Average latency $\times L$, for $K = 600$, and $M = 100$.

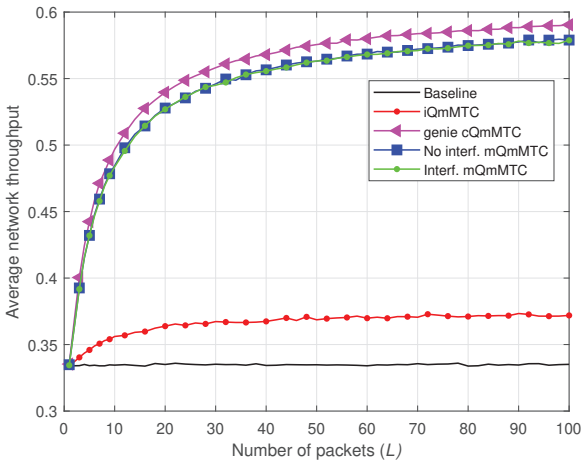


Fig. 7. Average network throughput $\times L$, for $K = 600$, and $M = 100$.

B. Random L_k and $P_a = 1$

In this subsection, we evaluate the scenario when the number of packets L_k sent by each device is random and follows a discrete uniform distribution as $L_k \sim \mathcal{U}[1, 10]$, while we still maintain $P_a = 1$ such that $K = K_a$. Figures 8 and 9 show, respectively, the graphs of latency and network throughput. One can note that the superiority of the results achieved by the proposed methods over the results achieved by the iQmMTC and baseline methods are preserved. Indeed, the curves in Figures 8 and 9 present practically the same shapes of the ones in Figures 2 and 3, respectively, but with a little performance degradation due to the reduction in the average number of transmitted packets, which limits the learning capability of the QL algorithm in seeking less congested pilots.

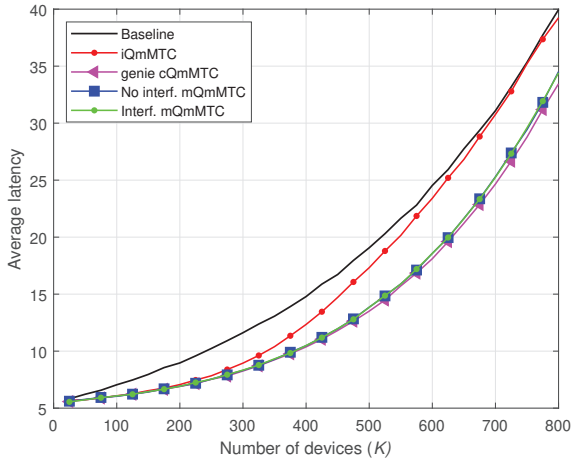


Fig. 8. Average latency $\times K$, for $L_k \sim \mathcal{U}(1{,}10)$, and $M = 100$.

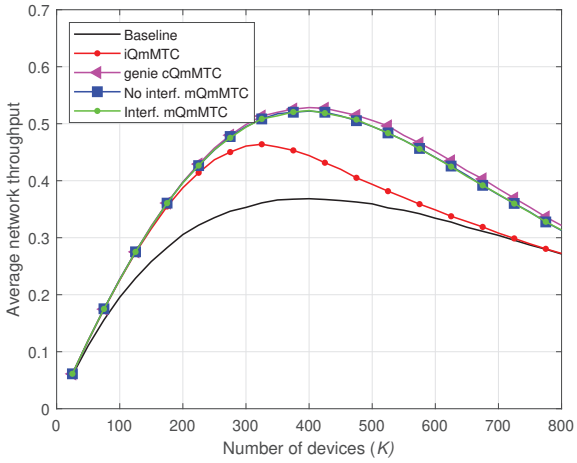


Fig. 9. Average network throughput $\times K$, for $L_k \sim \mathcal{U}(1{,}10)$, and $M = 100$.

C. Random L_k and $P_a = 0.1\%$

We consider in this subsection random number of packets sent by each device following $L_k \sim \mathcal{U}[1, 10]$, and a random number of devices being activated at each frame following a binomial distribution with activation probability of $P_a = 0.1\%$, such that $K_a \leq K$. The number of available RA pilots is also reduced to $\tau_p = 40$, in order to keep the simulation time not so long. The graphs presented in Fig. 10 and 11 are generated with 64000 Monte-Carlo realizations. Fig. 10 presents the average per-user throughput and Fig. 11 the average network throughput. The reward in the QL framework of the mQmMTC GF RA protocol is calculated using (19), while assuming that $\mathbb{E}[L_k] = 5.5$ is known at the devices' side. In terms of both performance metrics, our proposed mQmMTC protocol remain quite close to that of the ideal cQmMTC protocol, while always superior than that of Baseline and iQmMTC. While the per-user throughput of Baseline drops below 0.5 for $K \approx 2400$, this happens with $K \approx 2800$ for iQmMTC, with $K \approx 3100$ for mQmMTC, and with $K \approx 3400$ for the ideal cQmMTC. Similarly, the network throughput falling point in Fig. 11 occurs with $K \approx 2300$ for the Baseline, with with $K \approx 2400$ for iQmMTC, with $K \approx 2800$ for mQmMTC, and with $K \approx 3100$ for the ideal cQmMTC. These results corroborate the feasibility of the proposed mQmMTC GF RA protocol to address the challenges of massive machine-type communications in the framework of next generation massive multiple access systems.

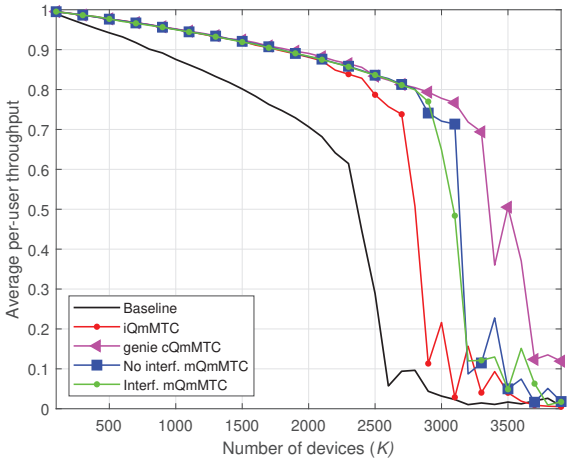


Fig. 10. Average per-user throughput $\times K$, for $L_k \sim \mathcal{U}(1{,}10)$, and $M = 100$.

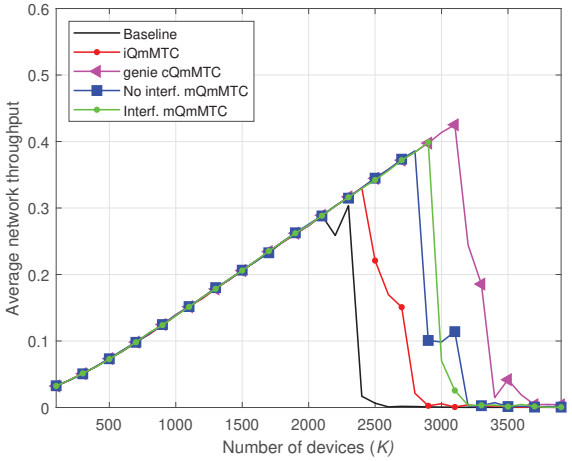


Fig. 11. Average network throughput $\times K$, for $L_k \sim \mathcal{U}(1{,}10)$, and $M = 100$.

V. Final Remarks

In this work, we have applied the collaborative, distributed and decentralized QL-based GF RA protocol to a massive MIMO scenario for pilot collision control, assuming realistic wireless propagation effects, such as multipath fading, shadowing, path loss, thermal noise and ICI. As the devices cannot know the exact number of pilot contenders without incurring excessive complexity and signaling overhead, our proposed approach takes advantage of the massive number of BS antennas to allow the devices to compute the QL rewards in a simplified way. We have also shown that our proposed approach is robust regarding the number of packets

to transmit, which can be as small as 10 or even random following a discrete uniform distribution, and regarding the number of active users, which can be randomly activated following a binomial distribution. Our proposed method is also robust regarding the number of antennas variation, and does not require more than ≈ 30 antennas at the BS to produce significantly improved performance, very close to the ideal (*genie*) cQmMTC protocol of [15].

CRediT authorship contribution statement

Felipe Augusto Dutra Bueno: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Investigation, Validation. **Alessandro Goedtel**: Conceptualization, Methodology, Writing – review & editing, Visualization, Investigation, Supervision. **Taufik Abrão**: Writing – review & editing, Visualization, Investigation, Supervision, Project administration, Funding acquisition. **José Carlos Marinello Filho**: Conceptualization, Methodology, Software, Data curation, Writing – review & editing, Visualization, Investigation, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 310681/2019-7.

References

[1] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive Access for 5G and Beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, 2021.

[2] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a Reality - What is Next?: Five Promising Research Directions for Antenna Arrays," *Digital Signal Processing*, vol. 94, pp. 3 – 20, 2019.

[3] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[4] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten Myths and One Critical Question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[5] M. Lee, Y. Kim, Y. Piao, and T.-J. Lee, "Recycling random access opportunities with secondary access class barring," *IEEE Transactions on Mobile Computing*, vol. 19, no. 9, pp. 2189–2201, 2020.

[6] E. Björnson, E. de Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A Random Access Protocol for Pilot Allocation in Crowded Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 2220–2234, April 2017.

[7] H. Han, L. Fang, W. Lu, K. Chi, W. Zhai, and J. Zhao, "A Novel Grant-Based Pilot Access Scheme for Crowded Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 11111–11115, 2021.

[8] J. C. Marinello and T. Abrão, "Collision Resolution Protocol via Soft Decision Retransmission Criterion," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4094–4097, 2019.

[9] H. Han, Y. Li, and X. Guo, "A Graph-Based Random Access Protocol for Crowded Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7348–7361, 2017.

[10] J. C. Marinello, T. Abrão, R. D. Souza, E. de Carvalho, and P. Popovski, "Achieving Fair Random Access Performance in Massive MIMO Crowded Machine-Type Networks," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 503–507, 2020.

[11] E. Casini, R. De Gaudenzi, and O. Del Rio Herrero, "Contention Resolution Diversity Slotted ALOHA (CRDSA): An Enhanced Random Access Scheme for Satellite Access Packet Networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1408–1419, 2007.

[12] G. Liva, "Graph-Based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 477–487, 2011.

[13] L. M. Bello, P. Mitchell, and D. Grace, "Application of Q-Learning for RACH Access to Support M2M Traffic over a Cellular Network," in *European Wireless 2014; 20th European Wireless Conference*, pp. 1–6, 2014.

[14] G. M. F. Silva and T. Abrão, "Throughput and latency in the distributed Q-learning random access mMTC networks," *Computer Networks*, vol. 206, p. 108787, 2022.

[15] S. K. Sharma and X. Wang, "Collaborative Distributed Q-Learning for RACH Congestion Minimization in Cellular IoT Networks," *IEEE Communications Letters*, vol. 23, no. 4, pp. 600–603, 2019.

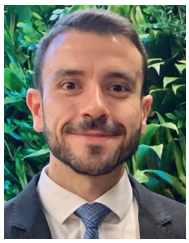
[16] 3GPP, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," Technical Report (TR) 25.996, 3rd Generation Partnership Project (3GPP), June 2018. Version 15.0.0.

[17] J. C. M. Filho, G. Brante, R. D. Souza, and T. Abrão, "Exploring the Non-Overlapping Visibility Regions in XL-MIMO Random Access and Scheduling," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6597–6610, 2022.

Felipe Augusto Dutra Bueno received his B.S. degree in Electrical Engineering from the University of Marilia, SP, Brazil, in Dec. 2017. He is currently pursuing an MSc. degree at the Federal University of Technology PR, Cornélio Procópio, Brazil from 2021. His current research interest includes signal processing and wireless communications systems, especially in MIMO, massive MIMO, massive machine-type communications, and random access protocols for crowded networks and B5G systems in general.

Alessandro Goedtel was born in Brazil, in 1972. He received the B.S. degree in electrical engineering from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 1997, the M.Sc. degree in industrial engineering from São Paulo State University, Presidente Prudente, Brazil, in 2003, and the Ph.D. degree in electrical engineering from the University of São Paulo, São Paulo, Brazil, in 2007. He is currently an Associate Professor with the Federal University of Technology—Paraná, Cornélio Procópio, Brazil. His research interests are within the fields of electrical machinery, intelligent systems, and power electronics.

Taufik Abrão (SM'12, SM-SBrT) received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from the Polytechnic School of the University of São Paulo, São Paulo, Brazil, in 1992, 1996, and 2001, respectively. Since March 1997, he has been with the Communications Group, Department of Electrical Engineering, Londrina State University, Paraná, Brazil, where he is currently an Associate Professor in Telecommunications and the Head of the Telecomm. & Signal Processing Lab. In 2018, he was with the Connectivity section, Aalborg University as a Guest Researcher. In 2012, he was an Academic Visitor with the Southampton Wireless Research Group, University of Southampton, Southampton, U.K. From 2007 to 2008, he was a Post-doctoral Researcher with the Polytechnic University of Catalonia, Barcelona, Spain. He has served as Associate Editor for the IEEE Systems Journal (2020), the IEEE Access (2016-2018), IEEE Communication Surveys & Tutorials (2013-2017), the AEUe-Elsevier (2020), the IET Signal Processing (2019), and JCIS-SBrT (2018-2020), and as Executive Editor of the ETT-Wiley (2016-2021) journal.



José Carlos Marinello Filho received his B.S. and M.S. degree in Electrical Engineering (the first with Summa Cum Laude) from Londrina State University, PR, Brazil, in Dec. 2012 and Sept. 2014, respectively, and his Ph.D. from Polytechnic School of São Paulo University, São Paulo, Brazil, in Aug. 2018. From 2015 to 2019, he was an Assistant Professor with Londrina State University, and, since 2019, he has been an Adjunct Professor with Federal University of Technology PR, Cornélio Procópio, Brazil. His research interests include signal processing and wireless communications, especially massive MIMO/XL-MIMO precoding/decoding techniques, acquisition of channel-state information, multicarrier modulation, cross-layer optimization of MIMO/OFDM systems, interference management in 5G, massive machine type communications, and random access protocols for crowded networks. He has been serving as Associate Editor for the SBrT Journal of Communications and Information Systems since 2019.

APPENDIX C – A NEURAL NETWORK-BASED RANDOM ACCESS PROTOCOL FOR CROWDED MASSIVE MIMO SYSTEMS

A Neural Network Based Random Access Protocol for Crowded Massive MIMO Systems

Felipe A. D. Bueno, Cézar F. Yamamura, Paulo R. Scalassara, Taufik Abrão, José Carlos Marinello.

***Abstract*—Fifth-generation and beyond networks are expected to serve large numbers of users equipments (UEs). Grant-based random access protocols are efficient when serving human users, which typically have large data volumes to transmit. The strongest user collision resolution (SUCRe) is the first protocol effectively taking advantage of the many antennas at the base station (BS) to improve connectivity performance. In this letter, we propose to replace the retransmission rule of SUCRe protocol with a neural network (NN) for identifying the strongest user, aiming to resolve the collisions in a decentralized way on the UEs' side. The NN is offline trained with different congestion levels of the system, aiming to obtain a single setup able to operate with different numbers of UEs. Our numerical results indicate that our proposed method obtains substantial connectivity performance improvements compared to other protocols without requiring additional complexity or overhead. Besides, the proposed approach is robust regarding variations in the number of BS antennas and transmission power.**

***Index Terms*—Random access protocol, grant-based protocols, massive MIMO, NNs, B5G.**

I. Introduction

Fifth-generation (5G) and beyond (B5G) mobile communication networks must be ready to provide reliable and enhanced Mobile Broadband (eMBB) communications to an ever-increasing number of devices [1]. Since time and frequency resources are scarce, pilot collisions during the random access (RA) stage will inevitably occur when the number of connected devices exceeds the number of available pilots, impairing the network as a whole. This problem is an important and challenging issue that future wireless networks must solve to provide reliable connections with the expected quality. Hence, the implementation of reliable and effective scheduling RA protocols is critical for the development of B5G networks.

There are multiple possible solutions for pilot collision in massive MIMO systems, such as the adoption of grant-free (GF) or grant-based (GB) protocols. One already validated approach is the strongest-user collision resolution (SUCRe) protocol [2]. This protocol is a 4-steps GB procedure that allows only the strongest contender to access the network resources. Although the SUCRe protocol can resolve about 90% of all collisions, it results in a high number of false-negative cases as shown in [3], since the SUCRe protocol does not resolve pilot collisions in which the strongest UE has a signal strength lower than 50% of the sum of the signal strengths of contending UEs.

Some works propose variations on the SUCRe protocol showing relatively good results. In [4], the UEs receive from the BS the precoded DL response to estimate the sum of the signal strengths of all contenders the UEs and also the information on which pilots remain idle after the first RA round and an access class barrier (ACB) factor for access control. Hence, some UEs that failed to be granted access on the first attempt can try to access the network resources through previously unused pilot signals. A similar protocol is proposed in [5] in which the BS broadcasts no ACB factor to UEs; instead, a graph-based interference cancellation scheme is applied to admit as many UEs as possible to the network. Both works show better results than the SUCRe protocol. However, both protocols introduce extra overhead to inform idle pilots to the UEs, increasing latency and harming the system's spectral efficiency. In [6], the access class barring with power control (ACBPC) RA protocol is introduced. The ACBPC protocol proposes a decentralized UL pilot transmit power control to be performed on the UEs' side. The results show a relative performance gain in comparison with the SUCRe protocol. Furthermore, the ACBPC protocol provides a fair access possibility to the UEs, independent of their distances to the BS. Another variation of the SUCRe protocol is a soft decision retransmission rule, softSUCRe, proposed in [7]. The softSUCRe rule differs from the SUCRe protocol as the UE decides to retransmit its pilot based on its probability of being the strongest user. The results are superior to the original SUCRe protocol. However, they present a small load of additional information for the UE to decide. Other works, such as [8], [9] and [10], also present different alternatives of collision resolution protocols. However, in [10], the SUCRe protocol is adapted for extra-large MIMO (XL-MIMO) systems, and [8] and [9] introduce extra overhead to the RA stage, making the system more complex as a whole.

Finally, in [3], a statistical-based GB RA protocol, aiming to resolve the collisions in a decentralized way at the UEs' side, is implemented with a bayesian classifier (BC) for identifying the strongest user and replacing the retransmission rule of SUCRe protocol. This protocol also shows results superior to the SUCRe protocol without the need for extra overhead. However, as a statistical approach, the BC method has a maximum accuracy limited by the distribution of the considered classes.

With the development of artificial intelligence, a new vision for solving the problems in communication systems especially using a neural network (NN) to pattern recognition has arised [11], [12]. This work proposes a decision-making methodology based on the Multi-layer Perceptron (MLP) NN applied to the SUCRe protocol empirical simulation data. The idea is to replace the retransmission rule of the original SUCRe with an MLP to indicate if the UE is the strongest contender or not, resolving thus the pilot collision in a decentralized way. The

performance results in the fraction of failed access attempts (FFAA) and the average number of access attempts (ANAA) indicate a significantly improved performance of our proposed method compared to other protocols available in literature.

The remainder of this paper is organized as follows. The system model is described in section II. The proposed NN-based RA protocol is explained in section III. The numerical results of the proposed protocol are presented, considering the ANAA and FFAA metrics, in Section IV. The main conclusions are offered in section V.

II. System model

Similarly as presented in [2], our model focus on a center hexagonal cell C_0, surrounded by 6 neighboring cells C_j with $j \in \{1, 2, ...6\}$. All cells present a BS located at their centers and equipped with M antennas to serve a set of UEs, through a time-division duplex (TDD) scheme, with time and frequency resources divided into coherence blocks of T channel uses. Furthermore, we represent the set of all UEs inside cell j by $\mathcal{U}_j$, and the subset of $\mathcal{U}_j$ of all active UEs by $\mathcal{A}_j \subset \mathcal{U}_j$. Also, we consider that inactive UEs will try to become active with probability $P_a \leq 1$. Therefore, even in cells with $|\mathcal{U}_j| \gg T$, it is possible to consider a scenario where $|\mathcal{A}_j| < T$. Such a scenario allows the BS to temporarily make orthogonal payload data pilot (PDP) signals available to all active UEs during payload data transmission by employing a grant-based RA protocol.

Let $\mathcal{K}_0 = \mathcal{U}_0 \backslash \mathcal{A}_0$ denote the set of inactive UEs with cardinality $K_0 = |\mathcal{K}_0|$ in cell C_0. The BS makes available a number τ_p of orthogonal RA pilot signals $\psi_1, \psi_2, ...\psi_{\tau_p} \in \mathbb{C}^{\tau_p}$ satisfying $||\psi_t||^2 = \tau_p$, $t \in \{1, 2, ..., \tau_p\}$. The τ_p available pilots are then shared by the K_0 inactive UEs. A particular UE that wants to become active firstly randomly chooses one pilot $\psi_{c(k)}$ out of the τ_p RA pilots signals available and then makes an access attempt by transmitting $\psi_{c(k)}$ with power $\rho_k > 0$, with $c(k) \in \{1, 2, ..., \tau_p\}$. The number of UEs that choose the pilot ψ_t and is denoted by $|\mathcal{S}_t|$, the cardinality of the set $\mathcal{S}_t = \{k : c(k) = t, \rho_k > 0\}$, which contains the of the UEs that choose the pilot t, and follows a binomial distribution [2]:

$$|\mathcal{S}_t| \sim \mathcal{B}\left(K_0, \frac{P_a}{\tau_p}\right). \tag{1}$$

The SUCRe protocol is divided into four steps as illustrated in the diagram show in figure 1:

i) Random Pilot Sequence: In the first step, the BS receives the signal $\mathbf{Y} \in \mathbb{C}^{M \times \tau_p}$ from the pilot sent by the UEs :

$$\mathbf{Y} = \sum_{k \in \mathcal{K}_0} \sqrt{\rho_k} \mathbf{h}_k \psi_{c(k)}^T + \mathbf{W} + \mathbf{N}, \tag{2}$$

where $\mathbf{N} \in \mathbb{C}^{M \times \tau_p}$ is the noise matrix of the signal received by the BS with each element following $\mathcal{CN}(0, \sigma^2)$, and $\mathbf{W} \in \mathbb{C}^{M \times \tau_p}$ is the interference signals received by the BS from the adjacent cells. The signal $\mathbf{Y}$ is then correlated with ψ_t by the BS:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Y} \frac{\psi_t^*}{||\psi_t||} = \textstyle\sum_{i \in \mathcal{S}_t} \sqrt{\rho_i} ||\psi_t|| \mathbf{h}_i + \mathbf{W} \frac{\psi_t^*}{||\psi_t||} + \mathbf{n}_t \\ &= \textstyle\sum_{i \in \mathcal{S}_t} \sqrt{\rho_i \tau_p} \mathbf{h}_i + \mathbf{W} \frac{\psi_t^*}{||\psi_t||} + \mathbf{n}_t, \end{aligned} \tag{3}$$

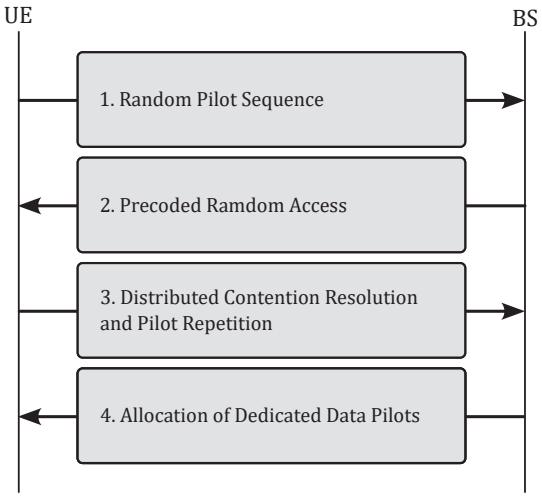


Fig. 1: SUCRe protocol diagram for massive MIMO.

where $\mathbf{n}_t = \mathbf{N} \frac{\psi_t^*}{||\psi_t||}$ is the effective noise, which follows a Random distribution $\mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$.

ii) Precoded Random Access: In the second step, the BS responds to all UEs who sent pilot signals with a precoded signal $\mathbf{V} \in \mathbb{C}^{M \times \tau_p}$:

$$\mathbf{V} = \sqrt{q} \sum_{t=1}^{\tau_p} \frac{\mathbf{y}_t^*}{||\mathbf{y}_t||} \psi_t^T, \tag{4}$$

in which q is the signal power available at this stage by the BS. The k-th UE then receives the signal $\mathbf{z}_k \in \mathbb{C}^{\tau_p}$:

$$\mathbf{z}_k^T = \mathbf{h}_k^T \mathbf{V} + \boldsymbol{\nu}_k^T + \boldsymbol{\eta}_k^T, \tag{5}$$

where $\boldsymbol{\nu}_k^T \in \mathbb{C}^{\tau_p}$ is intercellular interference (ICI) and $\boldsymbol{\eta}_k^T$ is the noise of the signal received by the UE, which follows a complex-normal random distribution $\mathcal{CN}(0, \sigma^2 \mathbf{I}_{\tau_p})$. Next, the UE correlates $\mathbf{z}_k$ with its chosen pilot ψ_t, resulting in

$$z_k = \mathbf{z}_k^T \frac{\psi_t^*}{||\psi_t||} = \sqrt{q\tau_p} \mathbf{h}_k^T \frac{\mathbf{y}_t^*}{||\mathbf{y}_t||} + \boldsymbol{\nu}_k^T \frac{\psi_t^*}{||\psi_t||} + \eta_k, \tag{6}$$

where $\eta_k \sim \mathcal{CN}(0, \sigma^2)$. Defining the sum of the signal strengths and interference received by the BS during the first step of the protocol for each pilot t, according to eq. (3) as α_t:

$$\alpha_t = \sum_{i \in \mathcal{S}_t} \rho_i \beta_i \tau_p + \omega_t. \tag{7}$$

Then, as proposed in [2], the value of α_t can be estimated by $\hat{\alpha}_{t,k}$:

$$\hat{\alpha}_{t,k} = \max\left(\left[\frac{\Gamma(M + \frac{1}{2})}{\Gamma(M)}\right]^2 \frac{q \rho_k \beta_k^2 \tau_p^2}{[\Re(z_k)]^2} - \sigma^2, \rho_k \beta_k \tau_p\right), \tag{8}$$

where $\Gamma(\cdot)$ is the gamma function, and $\Re(\cdot)$ returns the real part of a complex number.

iii) Distributed contention resolution and pilot repetition: In the third step, we assume that the k-th UE knows its own average channel gain β_k. Based on this value and the value $\hat{\alpha}_{t,k}$ received by each UE from the BS, the UE decides whether to retransmit the pilot signal or not. The main objective of the SUCRe protocol is to allow only one contending UE retransmit (the strongest one) to be connected to the network.

The pilot will be retransmitted when $\mathcal{R}_k$ is true and will not be retransmitted when $\mathcal{I}_k$ is true:

$$\mathcal{R}_k : \rho_k \beta_k \tau_p > \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{9}$$

$$\mathcal{I}_k : \rho_k \beta_k \tau_p \leq \frac{\hat{\alpha}_{t,k}}{2} + \epsilon_k, \tag{10}$$

in which ϵ_k is a bias parameter with a suitable value proposed in [2]. *iv) Allocation of Dedicated Data Pilots:* In the fourth step, all UEs that successfully retransmitted their pilots are granted access to exclusive network resources.

III. Neural Network Classifier

One of the most relevant features of artificial NNs is their capability of learning from the presentation of samples (patterns), which expresses the system behavior. Hence, after the network has learned the relationship between inputs and outputs, it can generalize solutions, meaning that the network can produce an output which is close to the expected (or desired) output of any given input values. In this section, we show how an NN classifier can be applied to the random access problem in crowded massive MIMO networks, under the strongest user criterion, to resolve pilot collisions. Aiming to resolve pilot collisions under the strongest UE retransmission rule on the UEs' side, we define two classes for the UEs: Z_0 represents the class of UEs which are not the strongest competitors for their chosen pilots, and Z_1 represents the class of the strongest UEs, forming the set $\mathcal{Z} = \{Z_0, Z_1\}$. We also define the state of each UE k as $\Omega_k \in \mathcal{Z}$. Our method works by seeking an approximation $\widehat{\Omega}_k$ of the class that truly represents the one of the k-th UE. This can be accomplished through an MLP NN by estimating a function $g(\cdot)$ that maps the input values $x_1 = \rho_k \beta_k \tau_p$ and $x_2 = \hat{\alpha}_{t,k}$ to the true state Ω_k of the k-th UE:

$$\widehat{\Omega}_k = \hat{g}(x_1, x_2), \tag{11}$$

where $\widehat{\Omega}_k$ is an estimator for the k-th UE's true class and $\hat{g}$ is an estimator for the function g.

The steps to realize this method are the following: *A. Database Acquisition*, *B. Preprocessing*, *C. Neural Network Training* and *D. Validation*

A. Database Acquisition

The first step is to acquire the NN training data. The database is generated from the simulation setup public shared by the authors of [2], where the values of $\hat{\alpha}_{t,k}$, β_k and their respective transmission classifications are collected, according to:

$$\mu_k = \begin{cases} 1 & \text{if } \Omega_k = Z_1 \\ 0 & \text{if } \Omega_k = Z_0. \end{cases} \tag{12}$$

B. Preprocessing

First, the input values are randomly mixed. Next, it is separated 80% for the NN training step and 20% for the NN validation. Then, the input values x_1 and x_2 are normalized according to:

$$\bar{x}_1 = \frac{\ln(x_1) - \min(\ln(x_1))}{\max(\ln(x_1)) - \min(\ln(x_1))}, \tag{13}$$

and,

$$\bar{x}_2 = \frac{\ln(x_2) - \min(\ln(x_2))}{\max(\ln(x_2)) - \min(\ln(x_2))}. \tag{14}$$

Finally, both input values are grouped into a vector $\bar{\mathbf{x}}_k$, where a bias term b is also appended:

$$\bar{\mathbf{x}}_k = [b \quad \bar{x}_1 \quad \bar{x}_2]^T, \tag{15}$$

where $(\cdot)^T$ is the transpose operation.

C. Neural Network Training

The training process of an NN consists of applying the required ordinated steps for tuning the synaptic weights and thresholds of its neurons, in order to generalize the solutions produced by its outputs. In the proposed method, the normalized data, $\bar{\mathbf{x}}_k$, and the desired output value μ_k associated with each $\bar{\mathbf{x}}_k$ sample are used as input data for training an MLP NN with okne hidden layer, as illustrated in Fig. 2. The MLP NN consists of a set of linear combinators, called

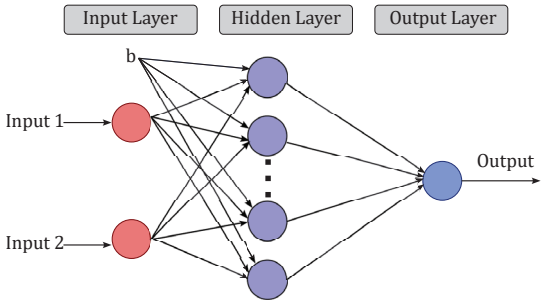


Fig. 2: MLP NN with one hidden layer

neurons, that control scalar product operations between input vectors and sometimes a bias term b with synaptic weights to generate a result by applying a given activation function. The bias term b accounts as an additional input term for the hidden layer's neurons. The training is carried out through the well-known backpropagation algorithm, which updates the set weight matrices $\mathbf{W_1} \in \mathbb{R}^{(r+1)\times t}$, which is the weight matrix between the input layer with r neurons and the hidden layer with t neurons and $\mathbf{W_2} \in \mathbb{R}^{t\times v}$, which is the weight matrix between the hidden layer with length t neurons and output layer with v neurons. Given a learning rate δ, the backpropagation algorithm, proceeds iteratively by minimizing the mean square error (MSE) function between the desired outputs $\mathbf{y}_k$ and the actual output $\hat{\mathbf{y}}_k$ at each i-th training epoch:

$$\text{MSE}_i = \frac{1}{2P} \sum_{k=1}^{P} (\mathbf{y}_k - \hat{\mathbf{y}}_k)^2, \tag{16}$$

where MSE_i is the MSE value at the i-th training epoch, and P is the total number of training samples. The training is considered complete when a given precision value,

$\epsilon = \text{MSE}_i - \text{MSE}_{i-1}$ is achieved. Once trained, the MLP NN can be used to estimate the function $g(\cdot)$ as:

$$\hat{\mathbf{y}}_k = s(\mathbf{W_2}^T \cdot s(\mathbf{W_1}^T \bar{\mathbf{x}}_k)), \tag{17}$$

where $s(\cdot)$ is the activation sigmoid function:

$$s(x) = \frac{1}{1+e^{-x}}. \tag{18}$$

Finally, the output $\hat{\mathbf{y}}_\mathbf{k}$ is associated to one of the output $\mathcal{Z}$ classes. Thus, deriving the estimator $\widehat{\Omega}_k$:

$$\widehat{\Omega}_k = \begin{cases} Z_0 & \text{if } \hat{\mathbf{y}}_k \leq 0.5 \\ Z_1 & \text{if } \hat{\mathbf{y}}_k > 0.5 \end{cases} \tag{19}$$

This work implements an MLP NN with $t = 5$ neurons in the hidden layer, $v = 1$ neurons in the output layer, and $r = 2$ neurons in the input layer, the bias term is set to $b = -1$. The learning rate is set to $\delta = 0.2$ and the mean square error precision of $\epsilon = 10^{-7}$.

D. Validation

After the training is complete, the MLP NN weights matrices $\mathbf{W_1}$ and $\mathbf{W_2}$ are validated according to (17). The purpose of the validation step is to verify whether the trained neural network meets the generalization criteria, by demonstrating its ability to generate results with a high degree of accuracy,*i.e*, above 50%, with input data that did not participate in the training phase and generate classification performances presented in the subsection IV-A.

IV. Numerical Results

In this section, we present the results obtained with the proposed method in terms of confusion matrices, ANAA, and FFAA .

We consider a system operating under the 5G sub-6 GHz band. The cell C_0 has a radius of $250\,\text{m}$ and a number of UEs K_0 varying between 100 and 30000 in increments of 500. The cells C_j with $j \in \{1, 2, ...6\}$ have the same radius and a fixed number of 10 active UEs each. The simulation parameters are set as M=100, ρ=1, q=1, τ_p=10, and transmission probability P_a=0.001. First, we provide results for the classification accuracy of the schemes, and then we present and compare the RA performance of the protocols.

A. Classification Performance

Tables I and II show the classification performance of the proposed MLP NN-based protocol in terms of confusion matrices. Table I presents the results for the scenario without ICI and table II results for the scenario with ICI. The bottom row shows the successful classification rates of each state, while the third column at the far right of the matrix shows the precision of predictions for each state belonging to $\widehat{\Omega}_k$. In other words, it presents the accuracy of the classifier given each output class $\widehat{\Omega}_k = Z_0$ or $\widehat{\Omega}_k = Z_1$. The far-right square at the bottom shows the overall accuracy of the classifier.

The results presented in Table I indicate that for the proposed MLP NN-based classifier, in a scenario without ICI, the classification success rate of UEs of class Z_0 is 99.1%, and the successful classification rate among Z_1 UEs is 76.5%, which is a number higher than the ones achieved by the SUCRe protocol, 41.7% and the ones achieved by the BC method, which is 74.8% [3]. Besides, the precision of Z_0 and Z_1 classifications are 98.3% and 85.8%, respectively. The overall accuracy of successful classifications is 97.5%, which is also higher than the values achieved by the SUCRe protocol and the BC method, which are respectively 96% and 97.3% in a scenario without ICI.

TABLE I: NN classifier w/o ICI

Predicted Class \ Actual Class	Z_0	Z_1	**total**
Z_0	**184560** 92.3%	**3226** 1.6%	98.3% 1.7%
Z_1	**1729** 0.9%	**10485** 5.2%	85.8% 14.2%
total	99.1% 0.9%	76.5% 23.5%	97.5% 2.5%

TABLE II: NN classifier with ICI

Predicted Class \ Actual Class	Z_0	Z_1	**total**
Z_0	**184068** 92.0%	**3599** 1.8%	98.1% 1.9%
Z_1	**1965** 1%	**10368** 5.2%	84.1% 15.9%
total	98.9% 1.1%	74.2% 25.8%	97.2% 2.8%

Table I presents the results for a scenario with ICI. The classification precision of Z_0 and Z_1 outputs are 98.1% and 84.1%, respectively. The percentages of correct predictions are 98.9% for class Z_0 UEs and 74.2% for class Z_1 UEs. The overall accuracy is 97.2%, higher than the results obtained with the SUCRe protocol and the BC method.

B. Connectivity Performance

Figures 3 and 4, respectively show the results of the metrics ANAA and FFAA in the following format: **a)** The baseline protocol of [2], which is an ALOHA-like protocol where pilot collisions are only handled by retransmission in later RA blocks, is shown in a black line. **b)** The red lines marked with "$\circ$" indicate the results of the original SUCRe protocol presented in [2]. **c)** The ciano lines marked with "$\triangleright$" indicate the results of the the ACBPC protocol of [6]. **d)** In the green lines marked with "x" are indicated the results obtained with the softSUCRe protocol of [7]. **e)** The magenta line with "$\diamond$" marker represents the results obtained with the BC method of [3]. **e)** Finaly, the blue lines with "$\square$" marker indicate

the results obtained when employing the MLP NN-based methodology proposed in this paper. The dotted line shows the results obtained for the cases without ICI. It is noteworthy the superiority of the proposed method. Compared to the BC method, for example, the proposed MLP NN-based method achieves a better performance when $K_0 \geq 25000$ UEs in the cases with ICI, as highlighted in Fig. 5, where K_0 varies from 25000 up to 40000.

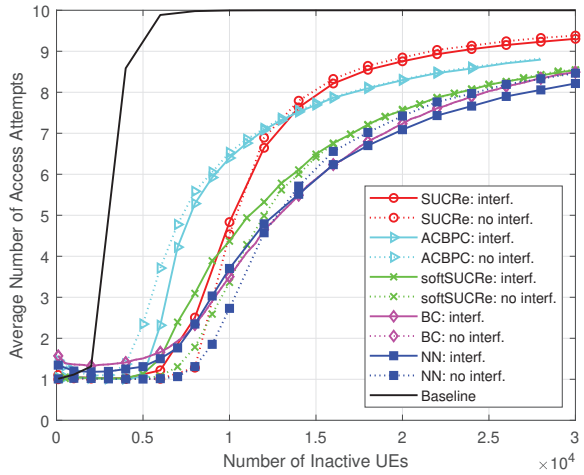


Fig. 3: ANAA $\times K_0$, for $M = 100$ and 0 dB of edge SNR

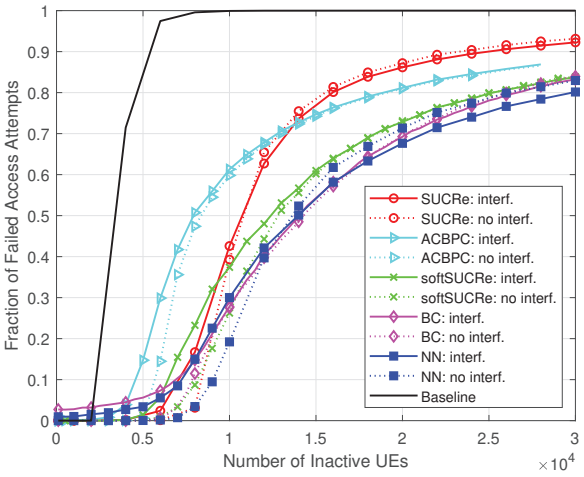


Fig. 4: FFAA $\times K_0$, for $M = 100$ and 0 dB of edge SNR

Fig. 6 shows the results of the ANAA and FFAA when the number of antennas varies from $M = 1$ to $M = 100$ in steps of 2. Even though the NN-based method is trained with $M = 100$ antennas, the proposed method shows robustness under the variation in the number of BS antennas M, where a number of $M \approx 50$ antennas reveals to be sufficient to provide superior results than the SUCRe and BC methods in both with and without ICI scenarios.

In Fig. 7 the edge SNR in dB (SNR_{dB}), defined as ρ/σ^2 and q/σ^2, with $\rho = q$, varies from -8 dB to $+8$ dB. The ANAA and FFAA are taken for a fixed number $M = 100$ antennas and $K_0 = 28000$ UEs. In both charts the performance

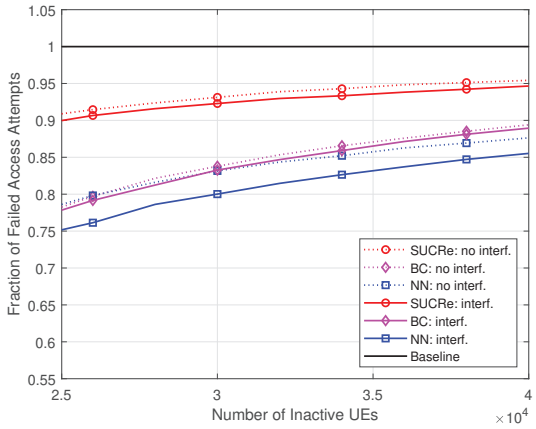


Fig. 5: FFAA $\times K_0$, for $M = 100$ and 0 dB of edge SNR

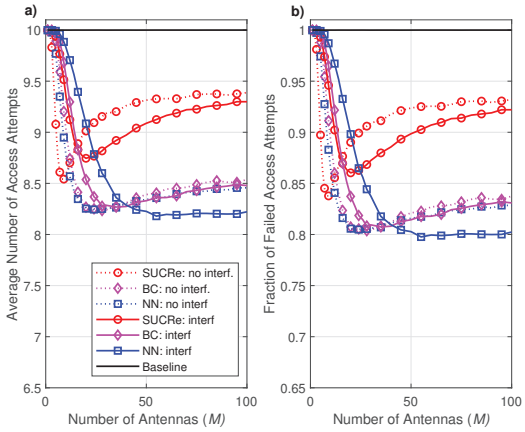


Fig. 6: Performance with M variation: **a)** ANAA, and **b)** FFAA.

of the NN-based method in both scenarios with and without ICI, is superior than the performance of the SUCRe protocol, in the whole considered edge SNR range. Compared to the BC method, the NN-based method presents a superior performance in a scenario with ICI from ≈ -6 dB to $\approx +6$ dB, and in the cases without ICI, from 0 dB to $\approx 6dB$. Consequently, the suggested approach demonstrates that it is a viable GB RA protocol option.

V. Conclusions

In this work, we have implemented a neural network based method aiming to optimize the SUCRe and other similar protocols. Based on our results, we conclude that the proposed neural network based method achieves significantly superior performance in comparison with SUCRe protocol, for both scenarios with and without inter-cellular interference and without the necessity of extra overhead. We have also shown how the proposed method is superior to other similar protocols, especially in overcrowded scenario, *i.e.*, $K_0 > 25000$. The robustness of the proposed method have also been assessed regarding the number of BS antennas and edge SNR, demonstrating that the proposed method is a promising GB RA protocol option.

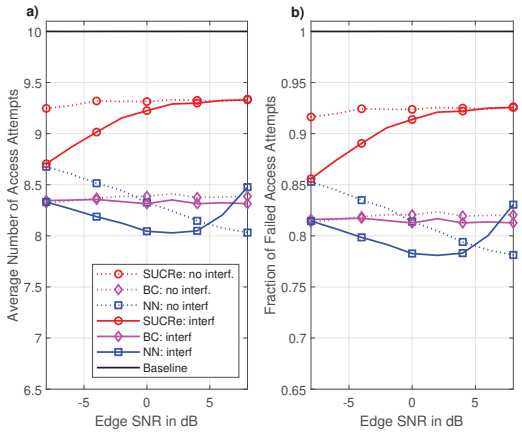


Fig. 7: Performance with edge SNR variation in dB: **a)** ANAA, and **b)** FFAA. With $M = 100$, and $K_0 = 28000$

REFERENCES

[1] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive Access for 5G and Beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, 2021.

[2] E. Björnson, E. de Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A Random Access Protocol for Pilot Allocation in Crowded Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 2220–2234, April 2017.

[3] Bueno, Felipe Augusto Dutra and Yamamura, Cézar Fumio and Goedtel, Alessandro and Marinello Filho, José Carlos, "A Random Access Protocol for Crowded Massive MIMO Systems Based on a Bayesian Classifier," *IEEE Wireless Communications Letters*, vol. 11, no. 11, pp. 2455–2459, 2022.

[4] H. Han, X. Guo, and Y. Li, "A High Throughput Pilot Allocation for M2M Communication in Crowded Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9572–9576, 2017.

[5] H. Han, Y. Li, and X. Guo, "A Graph-Based Random Access Protocol for Crowded Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7348–7361, 2017.

[6] J. C. Marinello, T. Abrão, R. D. Souza, E. de Carvalho, and P. Popovski, "Achieving Fair Random Access Performance in Massive MIMO Crowded Machine-Type Networks," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 503–507, 2020.

[7] J. C. Marinello and T. Abrão, "Collision Resolution Protocol via Soft Decision Retransmission Criterion," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4094–4097, 2019.

[8] H. Han, Y. Li, and X. Guo, "User Identity-Aided Pilot Access Scheme for Massive MIMO-IDMA System," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 6197–6201, 2019.

[9] H. Han, L. Fang, W. Lu, K. Chi, W. Zhai, and J. Zhao, "A Novel Grant-Based Pilot Access Scheme for Crowded Massive MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 11111–11115, 2021.

[10] O. S. Nishimura, J. C. Marinello, and T. Abrão, "A Grant-Based Random Access Protocol in Extra-Large Massive MIMO System," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2478–2482, 2020.

[11] J. Tao, J. Chen, J. Xing, S. Fu, and J. Xie, "Autoencoder neural network based intelligent hybrid beamforming design for mmwave massive mimo systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 1019–1030, 2020.

[12] P. Yang, J. Zhu, Y. Xiao, and Z. Chen, "Antenna selection for mimo system based on pattern recognition," *Digital Communications and Networks*, vol. 5, no. 1, pp. 34–39, 2019. Artificial Intelligence for Future Wireless Communications and Networking.