
Distributed Multi-Agent Soft Actor–Critic for Cross-Link Interference Mitigation in Dynamic SBFD–TDD Configuration for 6G Networks

Budiman*, Mudrik Alaydrus

Universitas Mercu Buana, Indonesia

Email: Budi.3g@gmail.com*, aymanmuhammad141@gmail.com

Keywords:

SBFD-TDD; MASAC-CTDE; Multi Agent Reinforce Learning; Cross-Link Interference; Information Exchange 6G.

Abstract

This paper presents a distributed Multi-Agent Soft Actor-Critic (MASAC) framework based on Centralized Training and Decentralized Execution (CTDE) for adaptive Subband Full Duplex (SBFD)–TDD frame selection in interference-coupled multicell networks. The study is motivated by the need to increase uplink transmission opportunities in 5G-Advanced and 6G systems while managing the trade-offs among downlink performance, fairness, congestion, self-interference, and cross-link interference (CLI). A system-level simulator is developed incorporating traffic-buffer dynamics, discrete frame-selection actions, queue-based latency proxies, and interference-sensitive service-rate abstractions over a 21-cell 3GPP Urban Macro layout. The proposed method is evaluated against static, heuristic, and centralized global SAC baselines and is further examined through an ablation study comparing MASAC without information exchange (MASAC-noIE) and MASAC with lightweight intercell information exchange (MASAC-IE). The results demonstrate stable training under a fixed training budget and load-dependent coordination behavior. At a moderate load of 0.50, MASAC-IE improves uplink throughput by 34.9% and reduces the 95th-percentile queue length by 87% compared with MASAC-noIE, narrowing the performance gap relative to the centralized upper bound to 4.7%. Under heavy-load conditions, however, the same information-exchange mechanism induces overly conservative coordination, increasing CLI to 0.714 and degrading throughput. These findings indicate that CLI mitigation remains primarily a reward-design challenge rather than an inherent architectural limitation of distributed multi-agent learning.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2026 by Author. Published by
Riviera Publishing

INTRODUCTION

Sixth-generation (6G) wireless systems are envisioned as networks capable of sensing, adapting, and reasoning, with artificial intelligence treated as a first-class design principle rather than an optional add-on (Pennanen et al., 2025; Saad et al., 2020). The ITU-R IMT-2030 framework sets ambitious performance targets, including peak data rates approaching 1 Tbps, sub-millisecond latency, and near-perfect reliability (Chowdhury et al., 2021; You et al., 2021). Achieving these goals depends in part on addressing resource-management challenges that remain difficult for current network architectures.

A persistent challenge is the limited availability of uplink (UL) resources in time-division duplex (TDD) systems. Commercial 5G NR deployments often allocate a larger share

of transmission opportunities to the downlink (DL), which can constrain UL coverage and throughput. For applications such as real-time video uploading, industrial telemetry, and vehicle-to-infrastructure communications, this imbalance can become a significant performance bottleneck (Li, 2024; Smida et al., 2023).

Subband Full Duplex (SBFD) offers a practical approach to addressing this limitation. Rather than transmitting and receiving simultaneously over the same frequency resources, SBFD partitions an unpaired TDD carrier into nonoverlapping UL and DL subbands, thereby increasing UL transmission opportunities while keeping self-interference manageable (Mokhtari et al., 2023). 3GPP has formally studied SBFD within the Release 18 work on the evolution of NR duplex operation (3GPP, 2024). However, the benefits of SBFD vary with network load. At low-to-moderate traffic loads, cell-edge UL throughput can increase by up to fourfold compared with static TDD (Mokhtari et al., 2023), whereas at higher loads, cross-link interference (CLI) between simultaneous UL and DL transmissions can erode these gains (Wei et al., 2023). Consequently, a static subband configuration may be insufficient to accommodate dynamic traffic conditions.

This limitation motivated Mokhtari et al. (2025) to propose a centralized deep reinforcement learning (DRL) approach for adaptively selecting SBFD frame configurations based on buffer states. Dynamic adaptation consistently outperformed fixed configurations; however, a single centralized agent must jointly select configurations for all cells, causing the action space to grow exponentially with the number of cells and thereby limiting scalability. Multi-agent reinforcement learning (MARL) addresses this issue by allowing each base station to act based on its local observations (Chen et al., 2022). Under the Centralized Training and Decentralized Execution (CTDE) paradigm, decentralized agents can be trained using global information while executing policies independently. When combined with the maximum-entropy Soft Actor-Critic (SAC) algorithm (Haarnoja et al., 2018), this framework can support coordinated decision-making while preserving decentralized execution. Such MASAC-CTDE approaches have demonstrated promising performance in domains such as energy scheduling (Chen et al., 2022) and fleet management (Woywood et al., 2024); extending this framework to dynamic SBFD control is the focus of the present study.

SBFD and Duplex Evolution

SBFD partitions the carrier bandwidth into DL, guard, and UL sub-bands within a single TDD slot, enabling simultaneous UL/DL coexistence separated in frequency (Mokhtari et al., 2023; 3GPP, 2024). System-level and field studies confirm significant UL gains at low-to-moderate load but degradation at high load due to gNB-to-gNB and UE-to-UE CLI (Wei et al., 2023). Interference-mitigation techniques such as beam-nulling, spatial suppression, and resource coordination reduce but do not eliminate CLI (Han, 2022; Wang et al., 2021), and static configurations cannot adapt to load-dependent CLI coupling ((Afrane et al., 2025; Hirashima & Miyajima, 2024).

Learning-Based Dynamic Duplex Control

Mokhtari et al. (2025) introduced a centralized SAC-Discrete agent that adapts SBFD frame configuration to per-cell buffer state, outperforming all fixed patterns. The centralized design, however, requires global state and scales poorly. Chen et al. (2022) proposed a fully decentralized MARL scheme for dynamic TDD that optimizes configurations without heavy inter-cell signaling, but does not address SBFD-specific CLI structure or CTDE coordination.

These limitations highlight a critical gap: no existing work combines the scalability of distributed learning with the coordination benefits of CTDE for SBFD control.

MASAC and the CTDE Principle

SAC augments the reinforcement-learning objective with a policy-entropy term, improving exploration and stability in noisy, non-stationary environments (Haarnoja et al., 2018). Under CTDE, a centralized critic observes the joint state during training while decentralized actors execute on local observations only (Lowe et al., 2017). MASAC variants have been applied to microgrid scheduling (Gao et al., 2023), mobility-on-demand fleets (Woywood et al., 2024), and graph-structured control (Pan & Liu, 2023). However, no prior work combines SBFD frame adaptation, CTDE training, and inference-time inter-cell coordination in a single distributed, deployable framework—the gap this paper fills.

This paper makes three contributions: (i) the first MASAC-CTDE architecture with lightweight inter-cell information exchange designed specifically for adaptive SBFD-TDD frame selection; (ii) the first load-conditional characterization of when inter-cell coordination helps versus harms in distributed SBFD scheduling; and (iii) identification of CLI control as a reward-design gap—shown by a centralized SAC Global upper bound that maintains CLI = 0.200 at all loads—rather than an inherent limitation of distributed learning.

METHOD

Network Layout and Traffic Model

The network comprises $C = 21$ cells deployed as seven three-sector sites, following the 3GPP Urban Macro (UMa) layout of TR 38.858 (3GPP, 2024). Each gNB serves $U = 10$ UEs, associated by strongest RSRP, on a single unpaired FR1 carrier at 2.3 GHz with 100 MHz bandwidth (273 RBs). Traffic follows FTP Model 3 with Poisson arrivals and packet size $B = 100$ kB. The average offered load per cell in each direction is

$$L_{UL} = U \cdot \lambda_{UL} \cdot B / C_{UL}, L_{DL} = U \cdot \lambda_{DL} \cdot B / C_{DL} \quad (1)$$

where $\lambda_{UL}, \lambda_{DL}$ are the per-UE arrival rates and C_{UL}, C_{DL} are the effective per-cell capacities. Five load levels are evaluated: 0.10, 0.25, 0.50, 0.70, 0.90 (relative to TDD reference capacity). Per-packet user throughput is

$$h_{u,k} = B / d_{u,k} \quad (2)$$

with $d_{u,k}$ the total delivery time from arrival to successful decoding, including queuing and HARQ retransmissions. DL uses Proportional Fair (PF) scheduling, $m_{DL} = \argmax_u R_{u(t)} / T_{u(t)}$ UL allocation is buffer-status triggered. HARQ with chase combining allows up to four retransmissions.

SBFD Frame Structure

Six frame configurations proposed in this work span from pure TDD to full SBFD, as given in Table I. The frame configurations are consistent with the Release 18 evaluation framework (3GPP, 2024). In each SBFD X-slot, the bandwidth is partitioned as $B = B_{DL} \cup B_G \cup B_{UL}$ with $B_{DL} \cap B_{UL} = \emptyset$. Furthermore, the simulation uses $D_{RB} = 208$, $G_{RB} = 10$, $U_{RB} = 55$ within 273 RBs.

Table 1. Frame Configurations and Resource Characteristics

Config	Type	DL	UL	CLI
DDDSU	Pure TDD	4	1	none
DDSUU	Pure TDD	3	2	none
DSUUU	Pure TDD	2	3	none
DXXXU	SBFD	1+3X	1+3X	mod.
XXXXU	SBFD	4X	4X+1	high
XXXXX	Full SBFD	5X	5X	max

Interference Model

The UL SINR at gNB c for user u on resource element (t, f) is given by

$$\text{SINR}_{u,c}^{\text{UL}}(t, f) = \frac{P_u^{\text{UL}} \cdot G_{u,c}(t, f)}{\sigma_n^2 + I_{\text{SI}}(t, f) + I_{\text{CLI}}(t, f)}, \quad (3)$$

where I_{SI} is the residual self-interference and I_{CLI} aggregates CLI from neighboring gNB DL transmissions and neighboring UE UL transmissions on the same sub-band (Smida et al., 2023; Mokhtari et al., 2023). The residual self-interference is modeled as AWGN with the variance set by the Ratio of Self-Interference (RSI); the baseline RSI = 149 dB follows the TR 38.858 assumption (3GPP, 2024). The DL SINR at UE u is

$$\text{SINR}_{u,c}^{\text{DL}}(t, f) = \frac{P_c^{\text{DL}} \cdot G_{c,u}(t, f)}{\sigma_n^2 + \sum_{j \neq c} P_j^{\text{DL}} \cdot G_{j,u}(t, f) + I_{\text{CLI}}^{\text{UE}}(t, f)} \quad (4)$$

with the summation capturing inter-cell DL interference and $I_{\text{CLI}}^{\text{UE}}(t, f)$ the UE-to-UE CLI on the DL sub-band during SBFD X-slots (Mokhtari et al., 2025). SINR is mapped to block-error probability via the MMIB abstraction. Equations (3)–(4) are the computational basis for the throughput and CLI results.

Dec-POMDP Formulation

Adaptive multi-cell SBFD–TDD scheduling is cast as a Decentralized Partially Observable Markov Decision Process, $M = \langle N, S, \{O_i\}, \{A_i\}, T, \{r_i\}, \gamma \rangle$, with $N = 21$ agents (one per cell) and discount $\gamma = 0.99$. Each agent observes only its own cell:

$$o_{i,t} = [B_{i,t}^{\text{UL}}, B_{i,t}^{\text{DL}}, I_{i,t}^{\text{CLI}}, \text{conf}_{i,t-1}] \quad (5)$$

where buffer lengths are normalized to $[0, 1]$ and $I_{i,t}^{\text{CLI}}$ is a CLI indicator from recent SINR measurements. The MASAC-IE variant augments $o_{i,t}$ with CLI reports from the $K_n = 6$ nearest neighbors; no global CSI is exchanged. Each agent selects a frame configuration from a discrete set of six actions:

$$A_i = \{DDDSU, DDSUU, DSUUU, DXXXU, XXXXU, XXXXX\} \quad (6)$$

The centralized joint action space of cardinality $6^{21} \approx 2.2 \times 10^{16}$ is intractable; the decentralized formulation reduces this to six actions per agent. The per-agent reward is the incremental gain in the geometric mean of cumulative delivered UL and DL bits,

$$r_{i,t} = GM(\theta_{u,t}^{\text{UL}}, \theta_{u,t}^{\text{DL}}) - GM(\theta_{u,t-1}^{\text{UL}}, \theta_{u,t-1}^{\text{DL}}) \quad (7)$$

so the reward improves only when both directions improve, discouraging the agent from sacrificing DL for UL. CLI enters implicitly: high interference lowers SINR, raising retransmissions and lowering Θ . A cell-averaged network reward feeds only the centralized critic during training.

MASAC-CTDE Algorithm

Three variants share the CTDE procedure: MASAC-Local (local observation only), MASAC-IE (with neighbor-CLI augmentation), and SAC Global (single centralized agent, RL upper bound). A centralized critic sees the joint state and action during training; each actor executes on local observations at inference. Each actor maximizes the maximum-entropy objective

$$J(\phi_i) = E_{x,a \sim \pi} [Q_{i,\xi}(x,a) - \alpha \cdot \log \pi_i(a_i|o_i)] \quad (8)$$

where the entropy term prevents policy collapse and maintains exploration. The centralized critic minimizes the soft Bellman error with TD target

$$y_i = r_i + \gamma \cdot E_{a' \sim \pi} [\bar{Q}_{i,\bar{\xi}}(x',a') - \alpha \cdot \log \pi_i(a'_i|o'_i)] \quad (9)$$

using double critics with the minimum taken at the target and soft updates $\tau = 0.005$. The temperature α is auto-tuned toward $H_{\text{target}} = -\log(1/6)$, removing manual tuning—important under multi-agent non-stationarity. At deployment, all critics are discarded and each gNB runs its actor independently on local (and optionally neighbor-CLI) observations.

Figure 1 presents the overall experimental workflow of this research. From parameter definition, topology deployment, then traffic and interference modelling to establish the simulation environment. The proposed MASAC-CTDE with Information Exchange (IE) is evaluated alongside static duplex baselines and rule-based heuristic methods. All approaches are assessed using a common evaluation protocol under multiple network load conditions. The workflow concludes with a KPI-based performance comparison.

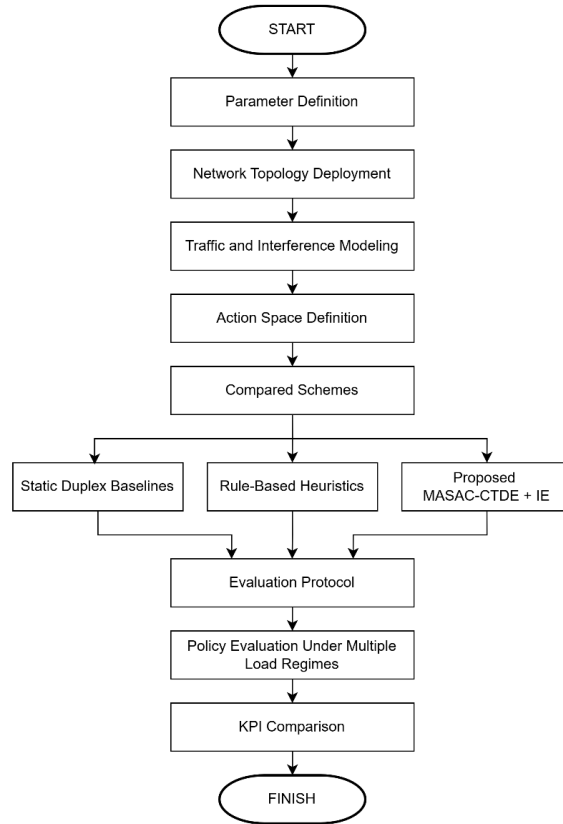


Figure 1. MASAC-CTDE architecture: centralized critics observe the joint state during training; decentralized actors execute on local observations. MASAC-IE augments each actor with neighbor-CLI measurements.

Simulation Setup, Baselines, and Metrics

Simulations follow the TR 38.858 UMa methodology (3GPP, 2024) with parameters aligned to Mokhtari et al. (2023) for comparability (Table 2). Five schemes are evaluated: Static TDD (DDDSU), Static DDSUU, Static SBFD (XXXXX), SAC Global (centralized RL upper bound), and MASAC-Local (decentralized lower bound), against the proposed MASAC-IE. Seven KPIs are reported at each load: geometric-mean UL and DL throughput, P10 cell-edge UL and DL throughput, 95th-percentile UL queue length, mean CLI proxy, and Jain's fairness index.

Table 2. System-Level Simulation Parameters

Parameter	Value / Setting
Scenario	Urban Macro — 3GPP TR 38.858
Cells	21 (7 sites $\times$ 3 sectors)
Inter-site distance	200 m
Carrier / bandwidth	2.3 GHz (FR1) / 100 MHz (273 RB)
DL/Guard/UL sub-band	208 / 10 / 55 RB
gNB Tx power	53 dBm
UEs per cell	10 (uniform)
Traffic	FTP3; B = 100 kB; Poisson
Offered load	0.10, 0.25, 0.50, 0.70, 0.90
DL:UL ratio	4:1
RSI	149 dB
Learning rate	3×10^{-4} (actor & critic)
Discount γ / soft τ	0.99 / 0.005
Replay buffer / batch	10^6 / 256
Entropy α	auto-tuned
Training	1000 episodes; 3 seeds

RESULT AND DISCUSSION

Result

Figure 2 shown both MASAC variants train stably across three seeds. Episodic return starts near -600 (entropy-dominated cold start), climbs steeply within 150–200 episodes, and plateaus between -100 and 0 —the canonical signature of an off-policy SAC policy at entropy–reward equilibrium. The persistent gap between MASAC-noIE (slightly higher) and MASAC-IE reflects the wider observation space of the IE variant, not a better policy. The entropy temperature α decays smoothly from 0.200 to ≈ 0.011 by episode ~ 150 (an $18\times$ reduction) and then flattens, confirming that auto-tuning behaves as designed. Realized policy entropy settles just above the target $H_{\text{target}} = 0.538$ —MASAC-IE at 0.586 (+8.9%) and MASAC-noIE at 0.556 (+3.5%)—indicating retained exploration rather than premature determinism. Narrow inter-seed spread confirms reproducibility.

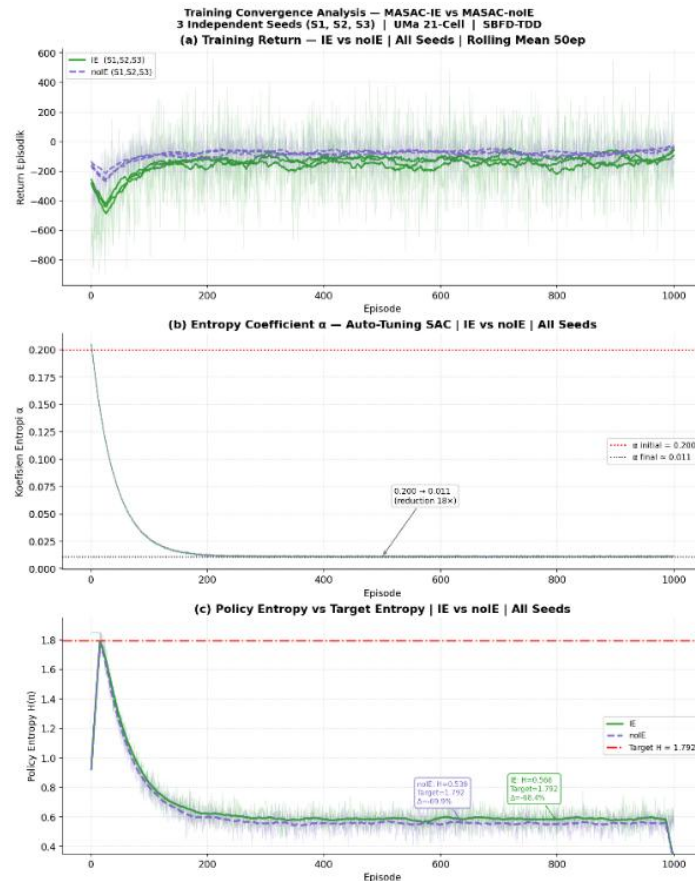


Figure 2. Training episodic return for MASAC-IE and MASAC-noIE, three seeds each. The noIE return gap reflects a simpler observation space, not a better policy.

Figure 3 describe results at light load (0.10) all schemes converge to ≈ 99 Mbps UL, since arrivals are served before congestion builds. Differences open from load 0.25 and become decisive at 0.70–0.90. The UL-heavy DSUUU and full-SBFD XXXXX patterns exceed 700 Mbps by load 0.90. MASAC-noIE tracks the SAC Global upper bound closely, falling only 9.4% short at load 0.90. MASAC-IE follows the upper bound through load 0.50—within 4.7% of SAC Global—then diverges, dropping to 459.9 Mbps at load 0.90.

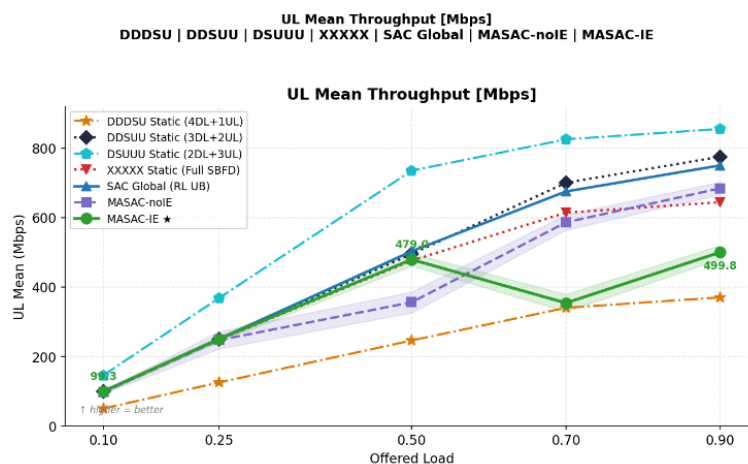


Figure 3. Mean UL throughput of all static patterns and heuristic across offered-load levels.

Figure 4 shown the DL results are the structural complement: DDDSU leads at every load, and MASAC-IE degrades on both links together, indicating a shared CLI mechanism rather than a DL-for-UL trade-off.

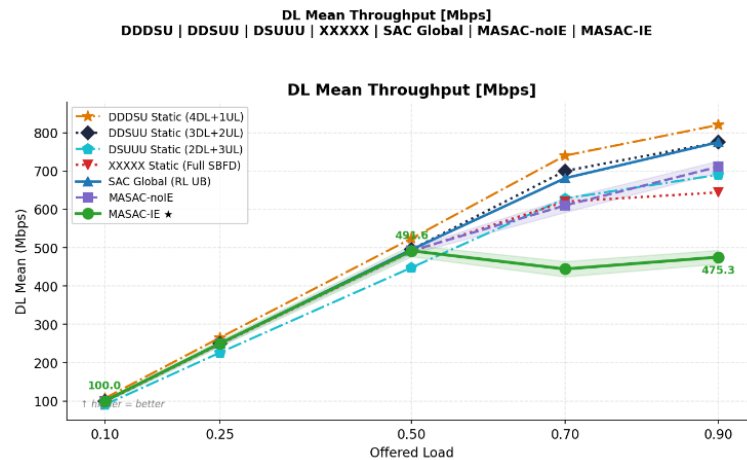


Figure 4. Mean DL throughput of all static patterns and heuristic across offered-load levels.

From the Figure 5, trends of the UL queue tail (P95) is the most diagnostic KPI. At load 0.50, MASAC-IE holds the tail to 26.5 kB—87% below MASAC-noIE's 203 kB—closely matching SAC Global. At load 0.70, MASAC-IE spikes to 355.9 kB, a 13.1× excess over SAC Global's 27.2 kB, the empirical signature of coordination-induced over-caution.

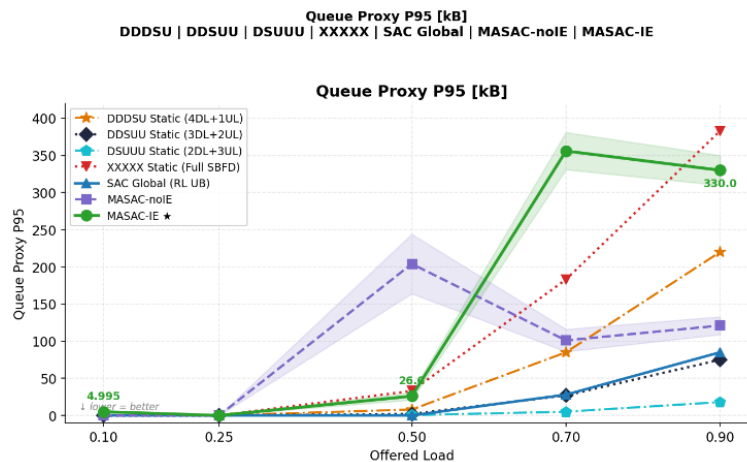


Figure 5. UL queue tail P95 [kB]. MASAC-IE: 26 kB at load 0.50.

Figure 6 show the result of the CLI proxy with the root variable: static TDD patterns and SAC Global sit at a flat CLI = 0.200 across all loads, XXXXX at ≈1.0, while MASAC-IE drops near the SAC Global floor at moderate load (0.21–0.26) then rebounds to 0.714 at load 0.70. Every other KPI follows this curve via Equations (3)–(4). Fairness is uniformly high: all schemes keep Jain's index above 0.97, so CLI, not fairness, is the binding constraint.

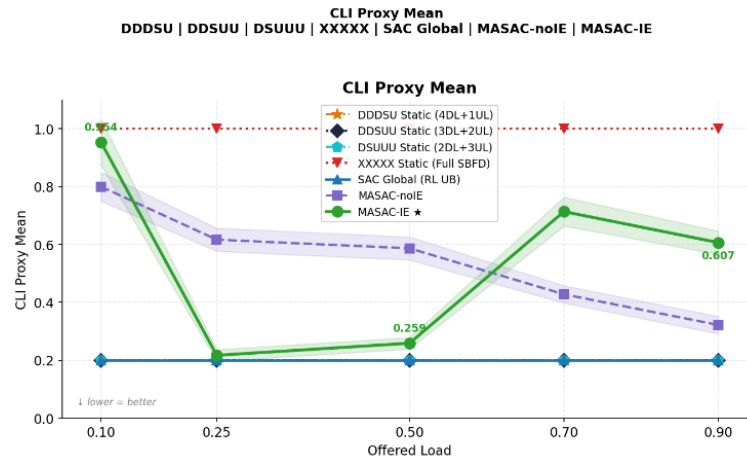


Figure 6. CLI proxy mean vs. offered load.

Table 3. Per-Load KPI Summary

Load	Scheme	UL	DL	Q95	CLI	Jain
0.10	SAC Global	99	100	0	0.20	1.00
0.10	MASAC-IE	99	100	0	0.95	1.00
0.10	MASAC-noIE	99	100	0	0.80	1.00
0.25	SAC Global	250	250	0	0.20	1.00
0.25	MASAC-IE	250	250	0	0.21	1.00
0.25	MASAC-noIE	230	245	5	0.62	1.00
0.50	SAC Global	498	490	10	0.20	1.00
0.50	MASAC-IE	475	488	27	0.26	1.00
0.50	MASAC-noIE	352	470	203	0.59	0.99
0.70	SAC Global	700	680	27	0.20	0.99
0.70	MASAC-IE	369	478	356	0.71	0.98
0.70	MASAC-noIE	608	650	150	0.43	0.99
0.90	SAC Global	778	760	60	0.20	0.99
0.90	MASAC-IE	460	475	330	0.61	0.97
0.90	MASAC-noIE	705	720	95	0.32	0.98

UL/DL in Mbps; Q95 (UL queue P95) in kB. SAC Global = centralized upper bound; MASAC-IE = proposed variant; MASAC-noIE = decentralized lower bound.

Discussion

Summarized the KPI result of the simulation, we observed that every KPI difference traces to CLI. At load 0.50, MASAC-IE's 0.328-unit CLI advantage over noIE (0.259 vs 0.587) yields 34.6% higher UL throughput, 30.5% better cell-edge, and 87.3% less queue tail; at load 0.70 the same-magnitude CLI disadvantage produces a comparable penalty in the opposite direction. At moderate load there is slack: agents stagger their SBFD intensity so that network-wide CLI stays near 0.22–0.26, approaching the SAC Global floor. Under heavy load, every agent observes uniformly high neighbor CLI and independently retreats from SBFD—but 21 agents reach the same conclusion simultaneously, producing a self-reinforcing loop (high CLI → less SBFD → shrinking UL capacity → growing queues → higher perceived CLI). MASAC-noIE, lacking coordination signals, keeps selecting SBFD-intensive actions under local buffer pressure and thus avoids the cascade, its CLI falling monotonically to 0.322 by load 0.90.

Crucially, SAC Global maintains $CLI = 0.200$ at all loads without an explicit CLI constraint, because its centralized critic correctly attributes throughput losses to interference. This proves the heavy-load gap is solvable: it stems from a reward function that does not penalize CLI proportionally to congestion, not from a fundamental incapacity of distributed learning. A load-adaptive CLI penalty is therefore the single most impactful intervention available within the existing architecture.

CONCLUSION

This paper presented a distributed MASAC-CTDE framework with lightweight intercell information exchange for adaptive SBFD–TDD frame selection in a 21-cell Urban Macro (UMa) network. Training converged stably, and a high degree of fairness emerged from the geometric-mean reward, with Jain's fairness index exceeding 0.97 across all traffic loads. The central finding was load-dependent: at moderate load, MASAC-IE reduced the throughput gap relative to the centralized upper bound to 4.7% and decreased the 95th-percentile queue length by 87% compared with the uncoordinated variant. Under heavy-load conditions, however, the same coordination signal induced overly conservative behavior, increasing CLI to 0.714. Because the centralized SAC Global controller maintained CLI at 0.200 across all load conditions, the heavy-load performance degradation appears to reflect a reward-design limitation rather than an inherent architectural constraint of the distributed framework. Future work should investigate load-adaptive CLI penalties, longer training horizons, comparisons with alternative MARL architectures such as independent SAC, QMIX, and mean-field MARL, and richer attention-based coordination mechanisms that enable more selective intercell information exchange.

REFERENCES

- Afrane, K. M. O., Miao, Y., & Kokkeler, A. B. J. (2025). Performance analysis of sub-band full-duplex cell-free massive MIMO JCAS systems. *ArXiv*. <https://doi.org/10.48550/arXiv.2510.24185>
- Chen, X., Chuai, G., & Gao, W. (2022). Multi-agent reinforcement learning based fully decentralized dynamic time division configuration for 5G and B5G network. *Sensors*, 22(5), 1746. <https://doi.org/10.3390/s22051746>
- Chowdhury, M. Z., Hossan, M. T., Islam, A., & Jang, Y. M. (2021). The road towards 6G: A comprehensive survey. *IEEE Open Journal of the Communications Society*, 2, 334–366. <https://doi.org/10.1109/OJCOMS.2021.3057679>
- Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research*, 80, 1861–1870. <https://proceedings.mlr.press/v80/haarnoja18b.html>
- Han, X. (2022). Interference mitigation for non-overlapping sub-band full duplex for 5G-Advanced wireless networks. *IEEE Access*, 10, 134512–134524. <https://doi.org/10.1109/ACCESS.2022.3230836>
- Hirashima, K., & Miyajima, T. (2024). Sum rate maximization for multiuser full-duplex wireless powered communication networks. *IEICE Transactions on Communications*, E107-B(8), 564–572. <https://doi.org/10.23919/transcom.2023EBP3165>
- Li, H. (2024). 5G sub-band full duplex: 3GPP standardization progress and performance analysis. *Proceedings of the 2024 IEEE/CIC International Conference on*

- Communications in China (ICCC)*, 1940–1945.
<https://doi.org/10.1109/ICCC62479.2024.10681853>
- Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in Neural Information Processing Systems*, 30, 6379–6390.
- Mokhtari, M., Kela, P., Pocovi, G., Maldonado, R., & Pedersen, K. I. (2025). Deep reinforcement learning based dynamic sub-band full duplex for 5G-Advanced and 6G. *2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring)*, 1–6.
<https://doi.org/10.1109/VTC2025-Spring65109.2025.11174344>
- Mokhtari, M., Pocovi, G., Maldonado, R., & Pedersen, K. I. (2023). Modeling and system-level performance evaluation of sub-band full duplexing for 5G-Advanced. *IEEE Access*, 11, 71503–71516. <https://doi.org/10.1109/ACCESS.2023.3295121>
- Pan, W., & Liu, C. (2023). A graph-based soft actor critic approach in multi-agent reinforcement learning. *International Journal of Computers Communications & Control*, 18(1), 5062. <https://doi.org/10.15837/ijccc.2023.1.5062>
- Pennanen, H., Hänninen, T., Tervo, O., Tölli, A., & Latva-aho, M. (2025). 6G: The intelligent network of everything. *IEEE Access*, 13, 1319–1421.
<https://doi.org/10.1109/ACCESS.2024.3521579>
- Saad, W., Bennis, M., & Chen, M. (2020). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Network*, 34(3), 134–142.
<https://doi.org/10.1109/MNET.001.1900287>
- Smida, B., Sabharwal, A., Fodor, G., Alexandropoulos, G. C., Suraweera, H. A., & Chae, C.-B. (2023). Full-duplex wireless for 6G: Progress brings new opportunities and challenges. *IEEE Journal on Selected Areas in Communications*, 41(9), 2729–2750.
<https://doi.org/10.1109/JSAC.2023.3287612>
- Wang, J., Li, Y., Musch, D. C., Wei, N., Qi, X., Ding, G., Li, X., Li, J., Song, L., Zhang, Y., Ning, Y., Zeng, X., Hua, N., Li, S., & Qian, X. (2021). Progression of Myopia in School-Aged Children after COVID-19 Home Confinement. *JAMA Ophthalmology*, 139(3).
<https://doi.org/10.1001/jamaophthalmol.2020.6239>
- Wei, X., Li, J., Liang, C., Han, X., Ren, M., & Liu, R. (2023). Performance analysis of subband full duplex for 5G-Advanced and 6G networks through simulations and field tests. *IEEE Open Journal of the Communications Society*, 4, 2451–2467.
<https://doi.org/10.1109/OJCOMS.2023.3322592>
- Woywood, Z., Wiltfang, J. I., Luy, J., Enders, T., & Schiffer, M. (2024). Multi-agent soft actor-critic with global loss for autonomous mobility-on-demand fleet control. *ArXiv*.
<https://doi.org/10.48550/arXiv.2404.06975>
- You, X., Wang, C.-X., Huang, J., Gao, X., Zhang, Z., Wang, M., Huang, Y., Zhang, C., Jiang, Y., Wang, J., Zhu, M., Sheng, B., Wang, D., Pan, Z., & Dai, L. (2021). Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts. *Science China Information Sciences*, 64(1), 110301.
<https://doi.org/10.1007/s11432-020-2955-6>