

# QoS Externalities in Network Sharing: A Congestion-Aware Nash Bargaining Framework

Maoquan Ni<sup>\*</sup>, Paolo Castagno<sup>†</sup>, Matteo Sereno<sup>†</sup>, Daniela Renga<sup>\*</sup>, Michela Meo<sup>\*</sup>  
<sup>\*</sup>Politecnico di Torino, Italy, <sup>†</sup>Università di Torino, Italy

**Abstract**—Network sharing (NS) between co-located mobile network operators can improve energy efficiency by consolidating traffic and allowing one base station (BS) to enter sleep mode. However, traffic aggregation increases congestion and may degrade quality of service (QoS), making the benefit of cooperation strongly load-dependent. This paper proposes a congestion-aware economic framework for NS that models QoS degradation as an externality and internalizes it through bargaining-based compensation. We combine a queueing-theoretic QoS model with an energy-aware BS power model to define operators' relative utilities with respect to a no-sharing baseline. A weighted Nash bargaining mechanism then determines the monetary transfer and allocates the cooperation surplus according to market share, activating NS only when cooperation is individually rational for both operators and yields non-negative total surplus. Using real mobile traffic traces and realistic BS power parameters, we show that the proposed QoS-aware rule yields a cooperative region close to, but more conservative than, a conventional hard headroom rule. The bargaining solution produces state-dependent transfer prices that increase with congestion pressure, while preserving host-side QoS and delivering measurable energy-efficiency gains. These results show that pricing QoS externalities provides a practical and economically meaningful basis for sustainable NS cooperation.

**Index Terms**—Network sharing, quality of service, congestion externalities, Nash bargaining, energy efficiency, queueing model

## I. MOTIVATION AND RELATED WORK

Mobile data traffic is projected to more than double by 2031, driven by the continued expansion of 5G services [1]. To satisfy peak-hour demand, radio access networks (RANs) are provisioned with substantial capacity headroom. Consequently, base stations (BSs) are often underutilized during off-peak periods while still consuming considerable energy, resulting in inefficient resource use and unnecessary emissions.

Because the RAN accounts for the largest share of mobile network energy consumption, improving its energy efficiency has become a strategic priority for mobile network operators (MNOs). At the same time, traffic demand is becoming increasingly dynamic across time and space due to emerging services and user mobility, making static dimensioning rules less effective. These trends motivate adaptive and economically grounded mechanisms that respond to load fluctuations while improving infrastructure utilization.

Network sharing (NS) (Fig. 1) has emerged as an effective approach to reduce deployment and operational costs while

improving energy efficiency through shared infrastructure and dynamic resource utilization [2]–[4]. By consolidating traffic across co-located operators, one BS can temporarily enter sleep mode while the other serves the combined demand, yielding potential energy savings. In 3GPP terminology, the active operator is the Master Operator (MOP), while the offloading operator is the Participating Operator (POP) [5].

However, the benefits of NS are strongly load-dependent. Consolidating traffic onto fewer active BSs increases utilization at the host BS and may degrade quality of service (QoS) through congestion, especially under high load. When utilization is low, the delay impact is limited; near saturation, even a modest traffic increase can sharply raise queueing delay and the probability of delay violations. Queueing-theoretic models are therefore widely used to evaluate delay and guide BS sleeping strategies [6].

The literature on energy-efficient BS sleeping and traffic consolidation is extensive. Measurements from operational networks show that cell switch-off with traffic offloading can deliver significant energy savings when supported by explicit QoS safeguards [7]. Analytical studies further characterize the energy–delay tradeoff and derive sleeping strategies that balance energy reduction against delay performance [6]. In most of these works, however, QoS is treated primarily as a technical constraint rather than an economic externality.

This creates an incentive asymmetry in NS: the POP benefits from switching off its BS, whereas the MOP bears the congestion risk and potential QoS degradation. Existing approaches often rely on fixed headroom constraints, allowing sharing only when host load remains below a predefined threshold. Although simple, such rules neither capture the marginal cost of congestion nor compensate the host operator, which may lead to inefficient or unstable cooperation.

Practical deployment of NS depends not only on technical feasibility but also on incentive alignment among competing MNOs. While economic and pricing models are well studied for 5G resource management [8], mechanisms tailored to NS remain limited. Related work on infrastructure sharing and network slicing considers revenue maximization, admission control, and dynamic pricing [9], [10], whereas congestion pricing has long been used for efficient wireless resource control [11]. Game-theoretic and coalition studies further show that cooperation depends on how surplus is allocated among operators.

Motivated by these gaps, this paper makes three contributions. First, we quantify load-dependent QoS degradation

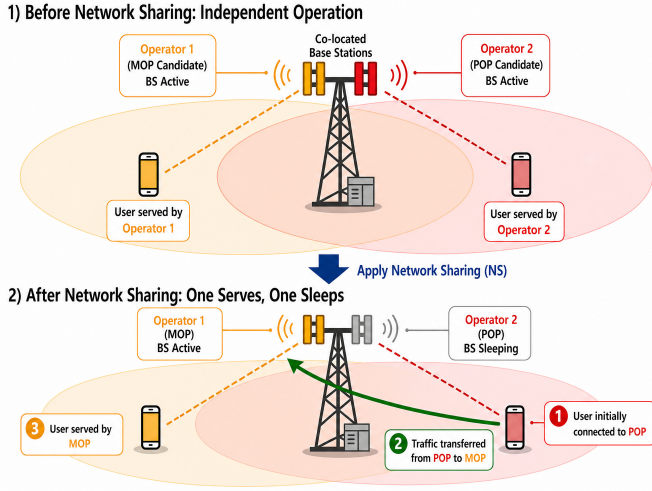


Fig. 1. Network sharing scenario: independent operation (top) vs. one-BS-active configuration after NS (bottom).

using a queueing-based model integrated with energy costs in an economic utility framework. Second, we propose a weighted Nash bargaining mechanism that internalizes congestion effects and allocates cooperation gains according to market share. Third, we characterize the tradeoff between energy savings and QoS degradation, identifying the cooperative region where sharing benefits both operators.

By jointly integrating congestion effects, energy consumption, and monetary transfers, the proposed framework provides a unified treatment of performance and incentives for sustainable NS cooperation.

## II. SYSTEM MODEL

### A. Network Sharing Scenario

We consider two MNOs operating co-located BSs over the same coverage area. Operator  $m \in \{1, 2\}$  owns a BS with service capacity  $\mu_m$  and traffic demand  $\lambda_m(t)$ , where time dependence is omitted when unnecessary.

At each decision interval, if the combined demand can be served by one BS, network sharing (NS) is activated: one BS remains active and serves the aggregate traffic, while the other enters sleep mode to reduce energy consumption. We denote the active operator as the host ( $h$ ), corresponding to the Master Operator (MOP) in 3GPP terminology, and the offloading operator as the guest ( $g$ ), i.e., the Participating Operator (POP) [5]. The host serves both traffic streams and bears congestion risk, whereas the guest switches off its BS. In this work, the BS with larger service capacity is selected as host.

Operational decisions are taken in discrete time slots, assuming short-term traffic forecasts are available.

As benchmark, we adopt the conventional headroom-based feasibility rule in [2]. Let  $\rho_h = \lambda_h/\mu_h$  be the host standalone load. Its residual normalized capacity is

$$R = 1 - \rho_h. \quad (1)$$

The host reserves a margin  $\hat{\rho}_*$  to avoid operating near saturation, so only  $R - \hat{\rho}_*$  is available for guest traffic. Under full offloading, NS is feasible if

$$\rho_{g \rightarrow h} \leq R - \hat{\rho}_*, \quad (2)$$

where  $\rho_{g \rightarrow h} = \lambda_g/\mu_h$  is the guest-induced load at the host. This fixed-headroom rule is later compared with the proposed QoS-aware economic framework.

### B. QoS and Congestion Model

To capture the nonlinear delay effect of traffic aggregation, QoS degradation is modeled using the average sojourn time of an M/M/1 queue:

$$S(\lambda, \mu) = \frac{1}{\mu - \lambda}, \quad \lambda < \mu. \quad (3)$$

Let  $\Lambda = \lambda_h + \lambda_g$  be the aggregated arrival rate under NS. For operator  $m \in \{h, g\}$ , the normalized delay variation is defined as

$$\Delta_{S,m}(\Lambda, \lambda_m) = \frac{S(\Lambda, \mu_h) - S(\lambda_m, \mu_m)}{S(\hat{\lambda}_m, \mu_m)}, \quad (4)$$

where  $\hat{\lambda}_m = (1 - \hat{\rho}_*)\mu_m$  is the normalization reference load.

For the host,  $\Delta_{S,h} = 0$  when  $\lambda_g = 0$ . For the guest, heterogeneous capacities may create a structural offset unrelated to offloading. To isolate the incremental congestion effect, we define the guest-side baseline correction

$$S_g^{(0)} = (\mu_g - \hat{\lambda}_g) \left( \frac{1}{\mu_h - \lambda_h} - \frac{1}{\mu_g - \lambda_g} \right), \quad (5)$$

which removes the standalone performance gap between host and guest. The corrected guest congestion term is then

$$\Delta'_{S,g} = \max\{0, \Delta_{S,g} - S_g^{(0)}\}, \quad (6)$$

so that  $\Delta'_{S,g}$  captures only NS-induced QoS degradation.

We also evaluate delay reliability through the probability that sojourn time remains below threshold  $\tau$ :

$$Q(\mu, \lambda; \tau) = 1 - e^{-(\mu - \lambda)^+ \tau}. \quad (7)$$

High values of  $Q$  indicate sufficient capacity margin, whereas low values imply congestion risk. In what follows, we write  $Q_i \triangleq Q(\mu_i, \lambda_i; \tau)$  for  $i \in \{h, g\}$  to denote the per-operator delay-reliability factor.

### C. Base Station Power Model

BS power consumption follows the EARTH framework [12]. For an active LTE BS,

$$P_{BS}(\lambda, \mu) = N_{\text{TRX}} \frac{P_{\text{PA}}(\lambda, \mu) + P_{\text{RF}} + P_{\text{BB}}}{(1 - \sigma_{\text{DC}})(1 - \sigma_{\text{MS}})(1 - \sigma_{\text{cool}})}, \quad (8)$$

where  $N_{\text{TRX}}$  is the number of transceiver chains, and  $\sigma_{\text{DC}}$ ,  $\sigma_{\text{MS}}$ , and  $\sigma_{\text{cool}}$  model supply and cooling losses.

The PA consumption is

$$P_{\text{PA},m} = \frac{\rho_m P_{\text{max},m}}{\eta_{\text{PA},m}(1 - \sigma_{\text{feed},m})}, \quad (9)$$

where  $\rho_m = \lambda_m/\mu_m$  is BS utilization. In sleep mode, the BS consumes constant residual power  $P_{\text{sleep},m}$ .

#### D. Energy Efficiency under Network Sharing

To assess how effectively NS converts power into delay-compliant service, we combine load and delay-reliability factors. With  $\rho_i = \lambda_i/\mu_i$ , define

$$\tilde{T} \triangleq \sum_{i \in \{h,g\}} Q_i \rho_i, \quad E \triangleq \frac{\tilde{T}}{P_{BS}(\lambda, \mu)}, \quad (10)$$

so load is discounted by its delay-compliance probability and  $E$  measures delay-compliant traffic per unit power.

### III. GAME-THEORETIC FORMULATION AND NASH BARGAINING

We model the interaction between the two operators as a cooperative bargaining game in which NS is a candidate action at each time slot, combining the energy and congestion effects introduced in Section II. Operators negotiate a monetary transfer price  $\eta \geq 0$  per unit of offloaded traffic: if an agreement is reached, traffic is offloaded and both energy savings and QoS degradation are realized; otherwise, each operator serves its own traffic independently and no transfer occurs. Cooperation is activated only when both operators achieve non-negative utility gains, defining the cooperative region as the set of individually rational operating points.

#### A. Utility Functions

Utilities are defined relative to the no-sharing baseline, so they represent gains or losses from cooperation rather than absolute profits. This avoids the need to model operator revenues explicitly. Each utility captures: (i) monetary transfers, (ii) variations in energy-related operating cost, and (iii) QoS degradation due to congestion.

When acting as guest, an operator offloads all traffic and may switch its BS to sleep mode. Its utility is

$$u_g(\eta, \lambda_h, \lambda_g) = \kappa_P \Delta_P(\lambda_g, 0) - \rho_{g \rightarrow h} \eta - \kappa_S \Delta'_{S,g}, \quad (11)$$

where  $\rho_{g \rightarrow h} = \lambda_g/\mu_h$  is the normalized load imposed on the host, and

$$\Delta_P(\lambda_1, \lambda_0) \triangleq P_{BS}(\lambda_1) - P_{BS}(\lambda_0).$$

When acting as host, the operator receives payment but incurs additional energy and congestion cost:

$$u_h(\eta, \lambda_h, \Lambda) = \rho_{g \rightarrow h} \eta - \kappa_{add} \Delta_P(\Lambda, \lambda_h) - \kappa_S \Delta_{S,h}, \quad (12)$$

where  $\Lambda = \lambda_h + \lambda_g$  is the aggregated load.

#### B. Disagreement Point and Individual Rationality

If no agreement is reached, both operators remain in standalone mode. Since utilities are measured relative to this baseline, the disagreement point is

$$(d_g, d_h) = (0, 0).$$

A sharing agreement is acceptable only if

$$u_g \geq 0, \quad u_h \geq 0.$$

Otherwise, NS is not activated.

#### C. Weighted Nash Bargaining Solution

The transfer price is determined through weighted Nash bargaining:

$$\max_{\eta} u_g^\alpha u_h^{1-\alpha} \quad \text{s.t. } \eta \geq 0, u_g \geq 0, u_h \geq 0, \quad (13)$$

where  $\alpha \in (0, 1)$  is the guest bargaining weight, interpreted as relative market share. Equivalently, using logarithms,

$$\max_{\eta} \alpha \ln u_g + (1 - \alpha) \ln u_h. \quad (14)$$

Under standard regularity conditions, the solution is unique and Pareto efficient.

#### D. Optimal Transfer Price

Since both utilities are affine in  $\eta$ , the problem is one-dimensional and concave. The first-order condition is

$$(1 - \alpha)u_g = \alpha u_h. \quad (15)$$

Solving for  $\eta$  gives

$$\eta^* = \frac{1}{\rho_{g \rightarrow h}} \left[ (1 - \alpha)(\kappa_P \Delta_P(0, \lambda_g) - \kappa_S \Delta'_{S,g}) + \alpha(\kappa_{add} \Delta_P(\Lambda, \lambda_h) + \kappa_S \Delta_{S,h}) \right]. \quad (16)$$

#### E. Total Cooperation Surplus

The total surplus is

$$\begin{aligned} \mathcal{S} &\triangleq u_g + u_h \\ &= \kappa_P \Delta_P(0, \lambda_g) - \kappa_{add} \Delta_P(\Lambda, \lambda_h) - \kappa_S \Delta'_{S,g} - \kappa_S \Delta_{S,h}. \end{aligned} \quad (17)$$

Since the transfer price  $\eta$  cancels out, it only redistributes value and does not affect the total surplus. Thus, cooperation is beneficial when  $\mathcal{S} > 0$ , while  $\mathcal{S} < 0$  implies that sharing should not occur. For  $\mathcal{S} \geq 0$ , admissible prices must also satisfy the individual rationality constraints.

#### F. Bargaining Properties

The weighted Nash solution satisfies Pareto efficiency, individual rationality, affine invariance, and proportional fairness. Substituting  $\eta^*$  into the utilities yields

$$u_g(\eta^*) = \alpha \mathcal{S}, \quad (18)$$

$$u_h(\eta^*) = (1 - \alpha) \mathcal{S}, \quad (19)$$

showing that efficiency is determined by  $\mathcal{S}$ , while the surplus allocation is governed by  $\alpha$ . In particular,  $\alpha = 1/2$  yields equal surplus division.

## IV. PERFORMANCE EVALUATION

#### A. Data and experimental setup

Our evaluation focuses on the city of Paris, France. We rely on real mobile traffic traces from the NetMob dataset, provided by a French mobile operator, to model BS traffic demand [13]. Traffic is reported as a 15-minute time series for each  $100 \times 100 \text{ m}^2$  geo-tile. For confidentiality reasons, the dataset is normalized by the data provider through a common

scaling factor, while remaining consistent and comparable across space and time.

We select an urban area obtained by aggregating  $n \times n$  contiguous geo-tiles. Within this area, we use the actual BS locations of the operator (denoted  $Op_1$ ) obtained from the public ANFR database [14]. To derive BS-level demand traces, we associate each geo-tile traffic time series to the geographically closest BS using a Euclidean distance criterion, thereby generating a Voronoi-like spatial partition. Let  $L = \{L_1, \dots, L_{n \times n}\}$  denote the set of tile traffic time series and  $\mathcal{B}_{(Op_1)} = \{B_{1,1}, \dots, B_{1,m}\}$  the set of BSs of  $Op_1$ . Each tile  $k$  is assigned to

$$j^*(k) = \arg \min_{j=1, \dots, m} d(L_k, B_{1,j}), \quad (20)$$

where  $d(\cdot, \cdot)$  denotes the Euclidean distance. The traffic demand at each BS is obtained by aggregating the time series of all tiles assigned to it.

To emulate two co-located operators while preserving realistic temporal dynamics and spatial correlations, we construct a second operator ( $Op_2$ ) by time-splitting the same dataset: traffic from week 1 is assigned to  $Op_1$ , while traffic from week 2 is assigned to  $Op_2$ , both mapped onto the respective BS locations. This approach preserves realistic demand variability while ensuring independence between operators. Consequently, at each 15-minute time slot  $t$ , a co-located BS pair observes traffic demands  $(\lambda_1(t), \lambda_2(t))$ .

Finally, for each BS we determine the service capacity  $\mu$  based on empirical peak demand to reflect realistic network dimensioning with headroom. Let  $\lambda_{\max}$  be the maximum observed demand over the considered time horizon. We define

$$\mu \triangleq \frac{\lambda_{\max}}{0.9},$$

so that the highest observed traffic load corresponds to 90% the largest service capacity. This choice introduces an explicit overprovisioning margin consistent with common operational deployment practices and ensures that the system operates below saturation under observed peak conditions.

In addition to the traffic traces, the evaluation specifies the key parameters used in the utility, bargaining framework, and power consumption in Table I.

## B. Numerical result

The shaded area in Fig. 2 denotes the operating points where both parties obtain positive utility ( $u_h > 0$  and  $u_g > 0$ ), while the color of the circular markers represents the host-side congestion variation,  $C_{\text{host}} = \Delta_{S,h}(\Lambda, \lambda_h)$ . Unlike an energy-only criterion, this construction explicitly incorporates congestion externalities into the cooperation decision.

When QoS costs exceed the energy gains from sharing, the cooperative region contracts and excludes points that are headroom-feasible but economically inefficient due to excessive service degradation. The square markers indicate such borderline unstable points. This occurs because the headroom condition only captures whether the host can technically absorb the guest traffic, not whether sharing remains worthwhile

TABLE I  
EVALUATION PARAMETERS (TWO SETTINGS FOR HEADROOM SENSITIVITY) AND BS POWER-MODEL PARAMETERS (MACRO BS).

Parameter	Setting A	Setting B
<i>NS decision / utility parameters</i>		
Reserved margin $\hat{\rho}_*$	0.1	0.2
Power weight $\kappa_P$	0.15	0.15
Additional power weight $\kappa_{\text{add}}$	0.03	0.03
QoS weight $\kappa_S$	0.09	0.1
<i>Bargaining / QoS evaluation parameters</i>		
Bargaining weight $\alpha$		0.453
Delay threshold $\tau$		$6.72 \times 10^{-8}$
<i>LTE Macro BS power model parameters</i>		
Number of transceiver chains $N_{\text{TRX}}$		6
Max transmit power $P_{\text{max}}$ [W]		39.8
PA efficiency $\eta_{\text{PA}}$		0.311
RF power $P_{\text{RF}}$ [W]		13
BB power $P_{\text{BB}}$ [W]		29.5
DC-DC loss $\sigma_{\text{DC}}$		0.075
Cooling loss $\sigma_{\text{cool}}$		0.10
Mains supply loss $\sigma_{\text{MS}}$		0.09
Sleep power $P_{\text{sleep}}$ [W]		75

once QoS costs are included. At higher host utilization, the QoS penalty grows rapidly, so some points satisfying the residual-capacity condition may still generate too little surplus for one or both operators. From an operational viewpoint, this means that technical feasibility alone is not sufficient for cooperation, since some feasible offloading decisions may still cause excessive service degradation at the host. Overall, the QoS-aware rule is safer than a pure threshold rule near saturation and yields a more meaningful, robust, and practically relevant cooperation boundary.

The two subplots correspond to Setting A and Setting B in Table I. Raising the headroom target from  $\hat{\rho}_* = 0.1$  in Setting A to  $\hat{\rho}_* = 0.2$  in Setting B makes the dashed feasibility boundary more restrictive and reduces its overlap with the utility-positive cooperative region. As a result, more points are excluded as non-cooperative, even near the feasibility boundary, highlighting the sensitivity of cooperation outcomes to the host's provisioning policy.

The host can further reshape the cooperation boundary through the QoS weight  $\kappa_S$ . A larger  $\kappa_S$  in Setting B places more weight on service degradation and shifts the cooperative region inward, yielding a more conservative sharing policy. A smaller  $\kappa_S$  in Setting A relaxes the QoS penalty and enlarges the set of utility-positive operating points. This shows that the proposed framework not only produces a more operationally sound and economically meaningful sharing boundary, but also captures how host-side policy choices directly shape it.

To connect the cooperative region with the operators' economic outcomes, we focus on Setting B in the remainder of this section, using the parameters in Table I, and examine the utility space at the feasible operating point.

Fig. 3 shows the feasible utility region in the  $(u_g, u_h)$  plane for  $R = 0.8$ . The region is convex, and individually rational outcomes are those in the non-negative quadrant relative to

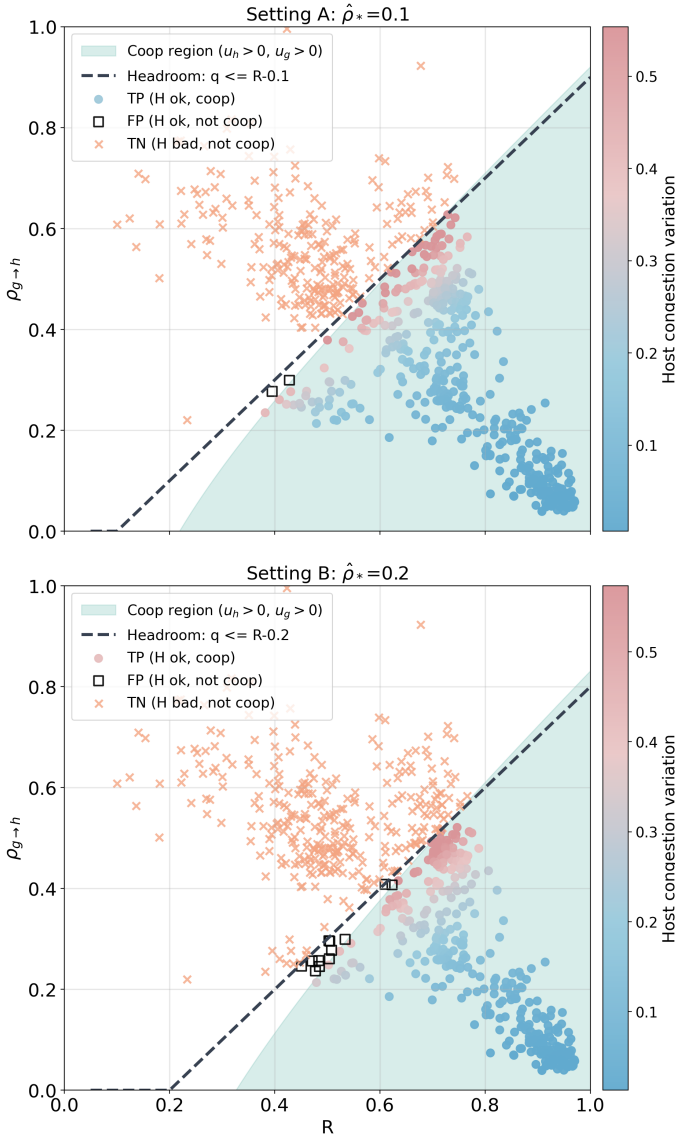


Fig. 2. The cooperative region produced by the proposed utility-based decision rule under two headroom settings,  $\hat{\rho}_* = 0.1$  (Setting A) and  $\hat{\rho}_* = 0.2$  (Setting B), compared with the hard headroom constraints.

the disagreement point  $(0, 0)$ . The Weighted Nash Bargaining Solution (WNBS) lies on the Pareto frontier, where any gain for one operator implies a loss for the other.

This highlights the role of bargaining in selecting a unique, Pareto-efficient, and weighted-fair outcome from the feasible network-sharing set. In particular, the WNBS determines the monetary transfer so that the cooperation surplus is split according to the bargaining weight  $\alpha$ , while preserving individual rationality for both guest and host.

Fig. 4 maps the negotiated monetary transfer over the cooperative region using contour lines. The contours reveal the nonlinear, state-dependent behavior of  $\eta^* \rho$ : their slope steepens as host utilization rises and as offloaded demand  $\rho$  increases. For fixed  $\rho$ , larger residual capacity  $R$  implies a lower transfer; for fixed  $R$ , increasing  $\rho$  raises the payment.

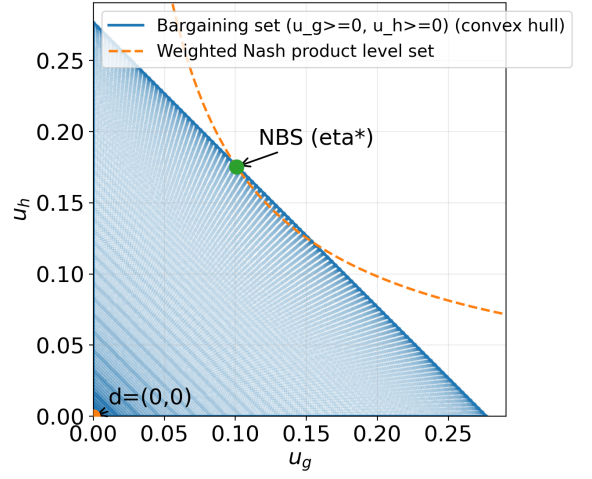


Fig. 3. Convex utility space for when the  $R=0.8$  and the WNBS point

This reflects the need to compensate the host for greater QoS degradation and higher marginal operating costs. Denser contours near the cooperative boundary further indicate rapidly growing externalities close to capacity limits, confirming that the bargaining-based pricing mechanism adapts the transfer to current network conditions.

Fig. 5 reports the QoS factor  $Q$  over three days for the two MNOs under standalone operation, i.e., without network sharing. The red curve shows the corresponding QoS factor when NS is active. As expected, NS causes some QoS degradation due to traffic aggregation, but the reduction remains limited, with  $Q$  never dropping below 95%.

Importantly,  $Q$  under NS is shown only during intervals in which sharing is beneficial, namely periods of low to medium utilization, typically at night and in the early morning. At higher utilization levels, the QoS term in the utility function discourages NS activation, prioritizing service quality over additional energy savings. These results indicate that NS can provide tangible energy savings while maintaining acceptable QoS when supported by an appropriate cooperation mechanism.

Having characterized the energy consumption profiles of the considered BS technologies, we now assess the system-level impact of enabling network sharing under the proposed decision rule. Fig. 6 validates the effectiveness of the utility-based mechanism in improving QoS-weighted energy efficiency. When cooperation is activated, the metric  $E(t) = (Q\rho)/P$  increases significantly, as traffic load is consolidated onto the host BS while the guest BS transitions to sleep mode. This consolidation reduces total system power consumption while maintaining acceptable QoS levels.

Averaged over all co-located BS pairs in the considered urban area and over a one-month evaluation window, the proposed cooperation policy yields a mean improvement of approximately 11% in QoS-weighted served traffic per kW relative to the no-sharing baseline. These results confirm that

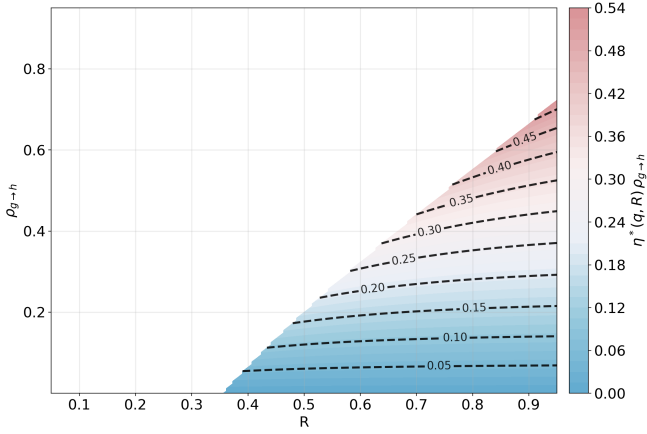


Fig. 4. Negotiated monetary transfer  $\eta$  over the cooperative region (contour lines)

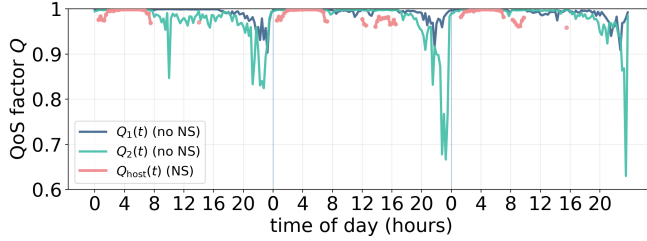


Fig. 5. QoS factor  $Q(t)$  over three days for both MNOs under standalone operation ( $Q_1$ ,  $Q_2$ ) versus network sharing ( $Q_{\text{host}}$ ).

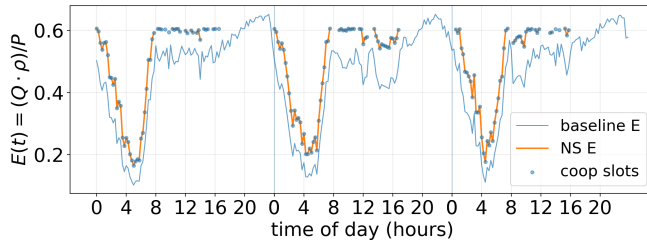


Fig. 6. QoS-weighted energy efficiency,  $E(t)$ , corresponding to the amount of delay compliant normalized traffic served per 1 kW of consumed power.

NS delivers intrinsic efficiency gains that extend beyond the purely monetary surplus captured in the bargaining framework, highlighting its potential as an operationally meaningful energy-optimization strategy.

## V. CONCLUSION

This paper presented a congestion-aware economic framework that provides a practical and interpretable basis for sustainable NS cooperation by jointly accounting for performance, energy savings, and incentive compatibility. The proposed method captures the QoS externalities caused by traffic aggregation and allocates the resulting cooperation surplus through a state-dependent transfer price. The results show

that the QoS-aware utility rule produces a cooperative region that is close to the conventional headroom-feasible region, but more selective and economically meaningful, excluding operating points that are capacity-feasible yet QoS-inefficient. The bargaining solution yields transfer prices that increase with load and congestion pressure. The trade-off evaluation further shows that the proposed NS policy preserves host-side QoS while delivering measurable improvements in energy efficiency. Future work will extend the model to multi-operator settings, uncertainty-aware traffic forecasts, and richer queuing abstractions such as M/G/1 and M/M/k/k to capture more general service-time distributions and finite-capacity effects.

## ACKNOWLEDGMENT

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”, focused project R4R).

## REFERENCES

- [1] Ericsson, “Mobile traffic forecast,” <https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast>, 2025, accessed: 2026-02-18.
- [2] D. Renga, M. Ni, M. A. Marsan, and M. Meo, “Network sharing to enable sustainable communications in the era of 5g and beyond,” in *ICC 2024 - IEEE Int. Conf. on Communications*, 2024, pp. 2840–2846.
- [3] E. J. Oughton, “Infrastructure sharing reduces the energy, emissions and costs of universal mobile 4g and 5g broadband,” *Telecommunications Policy*, vol. 49, no. 6, p. 102961, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0308596125000588>
- [4] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, “From network sharing to multi-tenancy: The 5g network slice broker,” *IEEE Communications Magazine*, vol. 54, no. 7, pp. 32–39, 2016.
- [5] 3GPP, “Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Network sharing; Concepts and requirements,” 3rd Generation Partnership Project (3GPP), Technical Specification TS 32.130, release 16.
- [6] P. Li, Y. Shen, F. Sahito, Z. Pan, and X. You, “BS sleeping strategy for energy-delay tradeoff in wireless-backhauling UDN,” *Science China Information Sciences*, vol. 62, no. 4, p. 42303, Feb. 2019.
- [7] D. Reiss, M. Catalan-Cid, D. Camps-Mur, and O. Sallent, “Quantifying the energy-saving and qos trade-off in traffic offloading for real 4g/5g scenarios,” in *IEEE International Conference on Communications (ICC) Workshops: GreenNet (BeGREEN)*, 2025, preprint / workshop paper.
- [8] N. C. Luong, P. Wang, D. Niyato, Y.-C. Liang, Z. Han, and F. Hou, “Applications of economic and pricing models for resource management in 5g wireless networks: A survey,” *Commun. Surveys Tuts.*, vol. 21, no. 4, p. 3298–3339, Oct. 2019. [Online]. Available: <https://doi.org/10.1109/COMST.2018.2870996>
- [9] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, and X. Costa-Perez, “Optimising 5g infrastructure markets: The business of network slicing,” in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, pp. 1–9.
- [10] Ö. U. Akgül, I. Malanchini, V. Suryaprakash, and A. Capone, “Dynamic resource allocation and pricing for shared radio access infrastructure,” in *IEEE International Conference on Communications (ICC)*, 2017.
- [11] B. Al-Manthari, N. Nasser, and H. Hassanein, “Congestion pricing in wireless cellular networks,” *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 358–381, 2011.
- [12] G. Auer, O. Blume, and V. Giannini, “Energy efficiency analysis of the reference systems, areas of improvements and target breakdown,” pp. 1–68, 12 2010.
- [13] O. E. Martínez-Durive, S. Mishra, C. Ziemlicki, S. Rubrichi, Z. Smoreda, and M. Fiore, “The netmob23 dataset: A high-resolution multi-region service-level mobile data traffic cartography,” 2023. [Online]. Available: <https://arxiv.org/abs/2305.06933>
- [14] Agence nationale des fréquences (ANFR). (2025) Observatoire 2g/3g/4g: Données sur les installations des réseaux mobiles télécoms.