Energy-Aware Load Shifting in SWIPT-Enabled IoT Networks

Biagio Boi,[†] Laura Finarelli,[*§] Falko Dressler,[§] Christian Esposito, [†] and Gianluca Rizzo[*‡]
{laura.finarelli,gianluca.rizzo}@hevs.ch, dressler@ccs-labs.org, {bboi,esposito}@unisa.it
[*] HES SO Valais [†] University of Salerno [‡] Università di Torino [§] TU Berlin

***Abstract*—The adoption of IoT is increasingly challenged by device energy constraints and growing environmental concerns, especially as device densities surge in future wireless networks. Energy Harvesting (EH) and Simultaneous Wireless Information and Power Transfer (SWIPT) offer promising solutions by enabling devices to recharge from ambient and wireless sources. However, efficiently managing network energy consumption without compromising connectivity remains unresolved due to SWIPT's nonlinear and non-monotonic energy dynamics. Existing load-shifting and traffic steering strategies assume linear load-energy relations, failing in the SWIPT context where both user and network energy dynamics fundamentally change. In this work, we bridge this gap by proposing a novel analytical framework for dynamic load shifting in SWIPT-enabled Radio Access Networks, leveraging stochastic geometry and realistic IoT energy models. We formulate an operator-centric optimization that accounts for delay-tolerant traffic and device duty cycling, targeting optimal load shifting strategies that ensure both communication and energy harvesting QoS. Numerical results suggest that our approach enables adaptive load distribution, strategically exploiting periods when network energy consumption is least sensitive to load and allows energy savings up to 15%. These findings highlight how integrating EH, SWIPT, and load management unlocks new avenues for scalable, energy-aware IoT networks, paving the way for sustainable 6G architectures.**

***Index Terms*—SWIPT, Energy Haversting, IoT**

I. Introduction

Despite its ubiquitous adoption, the IoT paradigm is still challenged by limited energy autonomy of devices, operational sustainability, and increasing environmental impact [1]. These issues are expected to intensify with the rise of massive machine-type communications (mMTC) and increasing device density, placing significant strain on future wireless networks [2], [3]. In this context, energy efficiency (EE) becomes essential, particularly in 6G heterogeneous networks where diverse users and devices coexist and share limited resources [4]. Energy Harvesting (EH) is thus emerging as a key strategy for sustainable IoT, enabling devices to convert ambient energy sources (e.g., RF, solar, kinetic) into usable power [5]–[7]. When combined with wireless power delivery [8], EH supports large-scale deployments, reduces reliance on finite batteries, and extends device lifetimes in remote or inaccessible environments.

An effective method to integrate EH into Radio Access Networks (RANs) is through Simultaneous Wireless Information and Power Transfer (SWIPT) networks [9], [10] (Figure 1), which enables concurrent data and power delivery. This dual functionality is particularly advantageous in battery-constrained or hard-to-reach deployments. By allowing devices to both communicate and recharge wirelessly, SWIPT enhances scalability, reduces operational costs, and supports energy-aware resource allocation strategies [11], [12]. However, ensuring energy-efficient operation in SWIPT networks while maintaining quality-of-service (QoS) for both energy and connectivity remains a major challenge [13], [14]. Unlike conventional cellular systems with predictable energy scaling, SWIPT networks exhibit non-monotonic energy dynamics due to the dual role of user devices. Indeed, all transmitting devices in a SWIPT RAN act not only as service consumers but also as distributed contributors to wireless power delivery. This shift undermines traditional energy management strategies, such as sleep modes or energy proportionality, which assume a quasi-linear relation between traffic load and energy consumption [13]. In particular, an increase in user density may lead to lower overall energy requirements and fewer active BSs needed to maintain QoS, within certain operating regimes [13], [15]. This phenomenon is due to emerging cooperative energy transfer among UEs, where active devices help power neighboring IoT nodes via passive EH.

Among the techniques with the largest potential for energy savings in traditional RANs, load shifting of delay-tolerant traffic via joint scheduling and traffic steering approaches, often empowered by AI and virtualization, has proven to enable significant improvements of both energy efficiency and network performance [16]–[18]. For example, [16] addressed joint scheduling and power control for offloading delay-tolerant traffic in dual-connectivity networks, showing significant energy savings through coordinated decisions. [17] advances policy-driven load balancing and traffic steering in O-RAN, leveraging two-tiered optimization to meet diverse traffic demands while reducing handovers and resource waste. [18] shows that adaptive on/off switching and data shift strategies can reduce macro BS power usage by up to 30%

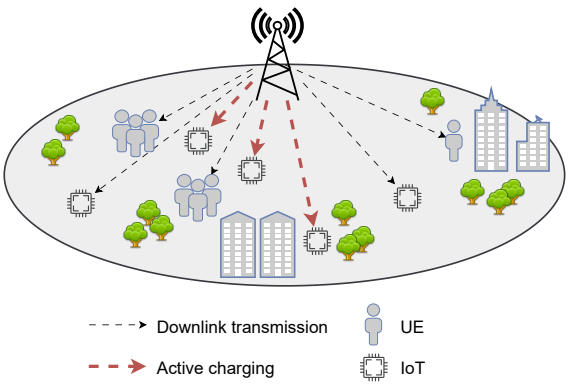

Fig. 1: Illustration of a SWIPT network where the base station provides connectivity for both BB and IoT UEs while actively charging IoT UEs.

per hour, especially for delay-tolerant services.

However, all these findings assume a quasi-linear and monotonic relationship between load and energy consumption, an assumption that does not apply to SWIPT networks, due to their inherently nonlinear and non-monotonic energy dynamics. Thus, it is unclear how can load-shifting strategies for delay-tolerant applications be redesigned or optimized in SWIPT-enabled RANs to maximize overall energy efficiency, while satisfying both connectivity and energy harvesting quality-of-service (QoS) constraints for heterogeneous IoT devices.

In this work, we apply recent results of stochastic geometry to investigate the potential of dynamic provisioning strategies based on load shifting in SWIPT networks. Specifically, our results can be summarized as follows:

- We propose an analytical model of SWIPT networks performance under delay tolerant IoT traffic, based on stochastic geometry and on an IoT energy model that accounts for the effects of duty cycling on user perceived performance.
- We formulate an optimization problem from the perspective of the Mobile Network Operator (MNO), to determine QoS-aware energy optimal load shifting strategies for IoT traffic.
- We characterize numerically the energy-optimal network provisioning strategies emerging from our framework. On a realistic daily traffic profile, we show that our framework is effectively able to shift the load towards those operating conditions in which the power consumed is the least sensitive to load.

Our results demonstrate the potential of combining EH, SWIPT, and DSM to create adaptive, energy-aware IoT architectures, laying the foundation for scalable and sustainable 6G wireless networks.

II. System Model

We consider a cellular network where BSs are distributed in space according to a homogeneous planar Poisson Point Process (PPP) Φ_b with intensity λ_b BSs per m^2. Similarly, UEs are distributed in space according to a homogeneous PPP $\Phi_u(t)$ with intensity of λ_u UEs per m^2. UEs can be broadband (BB) or IoT, though our analysis can easily be extended to account for other categories of users.

We consider the case in which IoT devices harvest the energy necessary for their operation from RF signals. Thus, every IoT device is capable of exploiting downlink signals from its serving BS to decode its intended information and charge its battery (*active charging*), as well as any signal transmitted from BSs and UEs for energy harvesting (*passive charging*).

Thus, each IoT device can use downlink signals from its serving base station to decode its intended information and to actively recharge its battery (*active charging*), and it can also harvest energy from any signals transmitted by base stations and user equipments in its vicinity (*passive charging*). Specifically, each IoT device is equipped with two separate receivers, one for information transfer and the other for energy harvesting. IoT devices commute between energy harvesting and information decoding via Time Splitting (TS) [19]. In such an operating mode, a fraction η (the *time split ratio, with* $0 \leq \eta \leq 1$) of the BS time dedicated to serving that device is devoted to active power transfer, while the remaining fraction $1-\eta$ of BS time dedicated to that user is devoted to receiving information. When an IoT device is not actively served by the BS to which it is associated, it is performing passive harvesting. We assume η takes the same value for all devices.

A. Service model

We assume BSs use a generalized processor-sharing (GPS) mechanism to divide BS time among all the connected devices. In the downlink, the GPS weights are 1 for IoT devices, and ω_d for BB UEs. To take into account the difference in QoS between IoT and BB, we define $\omega_d = \delta_d(1-\eta)$, where δ_d defines the ratio of QoS between the downlink process of the two categories of UEs. The time spent by the BS without transmitting is modeled as a user with a GPS weight β_d. As for uplink, in all configurations, the GPS weights are 1 for IoT UEs, δ_u for BB UEs, and β_u for the uplink BS time not assigned to any UE. We assume that BB UEs always have data to transmit and receive, and we denote the mean per-bit delay required as τ_u^0 for the uplink and τ_d^0 for the downlink. Instead, at each IoT device, we assume that the data to transmit arrives at a constant rate equal to T_u^0 bits per second. Similarly, at each BS, the data to be transmitted in the downlink to each IoT device arrives at a constant rate of T_d^0 bits per second.

We assume that, to optimize energy consumption (and thus autonomy), IoT UEs follow periodic cycles of activity and idle time, independently in downlink and uplink. Specifically, they transmit for a fraction ϕ_u of their time and receive data for a fraction ϕ_d of their time. This models IoT UEs whose activity is triggered by human presence, such as event detection systems.

To ensure fairness in the system, the values of ω_d and δ_u are set as the ratio between the traffic requirements of BB users and that of IoT UEs in active state in downlink and in uplink, respectively [20]. Thus, in general, the BS time dedicated to serving an IoT device is a fraction ϕ_d of the time that the BS might dedicate to serving it according to the GPS weight.

The BS utilization is the fraction of time during which it is busy serving all associated users. Specifically, if $S(x)$ denotes the BS serving a user located in x, the utilization in downlink is

$$U_d(S(x)) = \frac{\phi_d N_{iot}(S(x)) + \omega_d N_{bb}(S(x))}{\phi_d N_{iot}(S(x)) + \omega_d N_{bb}(S(x)) + \beta_d} \quad (1)$$

while in uplink, we have

$$U_u(S(x)) = \frac{\phi_u N_{iot}(S(x)) + \delta_u N_{bb}(S(x))}{\phi_u N_{iot}(S(x)) + \delta_u N_{bb}(S(x)) + \beta_u} \quad (2)$$

$N_{iot}(S(x))$ and $N_{bb}(S(x))$ are the number of IoT and BB UEs associated with $S(x)$, while $\phi_d N_{iot}(S(x))$ (respectively, $\phi_u N_{iot}(S(x))$) denote the mean number of IoT UEs in active state in downlink (resp. in uplink). Thus, given ω_d and δ_u, by tuning β_j, $j \in \{u, d\}$, the network controls the mean amount of service received by UEs for both communications and energy transfer, and the overall BS utilization, both in downlink and uplink.

We define the *ideal-per bit delay* τ_j with $j \in \{d, u\}$ as the per bit delay which a UE would perceive if the BS with which the UE is associated had utilization equal to one. Leveraging tools from stochastic geometry and following the proof in [20], we can define the ideal per-bit delay perceived by a BB UE at distance x from its nearest BS as:

$$\tau_d^{id}(x) = \frac{\phi_d N_{iot}(S(x)) + \omega_d N_{bb}(S(x))}{\omega_d C(x, P, G, I)}, \quad (3)$$

$$\tau_u^{id}(x) = \frac{\phi_u N_{iot}(S(x)) + \delta_u N_{bb}(S(x))}{\delta_u C(x, P_I, 1, 0)}. \quad (4)$$

We denote the capacity of a user located at a distance r from the BS by $C(r, P, G, I)$ bit/s per Hertz, where P is the BS transmit power, and I the total received interfering power. Using Shannon's capacity law, we model $C(r, P, G, I)$ as $C(r, P, G, I) = \frac{B}{\rho} \log_2 \left(1 + \frac{PGr^{-\alpha}}{N_0 + I(r,\rho)}\right)$, where α is the attenuation coefficient, N_0 the power spectral density of the additive white Gaussian noise, ρ the reuse factor. We assume BS antennas use beamforming, and we denote with G the beamforming gain and with L the side lobes attenuation.

B. Analytical Model

In this subsection, we present the main analytical results used in our optimization framework, based on recent stochastic geometry results. They are derived from [12], [13].

Theorem 1. *The mean ideal per-bit delays in downlink and uplink are given by:*

$$\bar{\tau}_d = H(\omega_d, \omega_d, C(r, P, G, \bar{I})) \quad (5)$$

$$\bar{\tau}_{d,I} = \bar{\tau}_d \frac{\omega_d}{(1-\eta)} \quad (6)$$

$$\bar{\tau}_u = H\left(\delta_u, \delta_u, C(r, P_I, 1, 0)\right) \quad (7)$$

$$\bar{\tau}_{u,I} = \delta_u \bar{\tau}_u \quad (8)$$

Where:

$$H(y, z, g(r)) = \int_0^\infty \frac{f(r, y) e^{-\lambda_b \pi r^2} \lambda_b 2\pi r}{z g(r)} dr. \quad (9)$$

with

$$f(r, y) = \lambda_u \left[y + \gamma\left(\phi_j - y\right)\right] \int_0^\infty \int_0^{2\pi} e^{-\lambda_b A(r,x,\theta)} x d\theta dx$$

$A(r, x, \theta)$ is given by $A(r, x, \theta) = \pi x^2 - \Bigg[r^2 \arccos\left(\frac{r + x\sin(\theta)}{d(r,x,\theta)}\right) + x^2 \arccos\left(\frac{x + r\sin(\theta)}{d(r,x,\theta)}\right) + - \frac{1}{2}\sqrt{[r^2 - (d(r,x,\theta) - x)^2][(d(r,x,\theta) + x)^2 - r^2]}\Bigg]$, and $d(r, x, \theta)$ is the euclidean distance between (x, θ) and $(0, -r)$ and $j \in \{d, u\}$. The interference term $\bar{I}$ in $C(r, P, G, \bar{I})$ is given by $\bar{I}(r, \rho) = \frac{P L_g \lambda_b 2\pi r^{2-\alpha}}{\rho(\alpha - 2)} \frac{\bar{\tau}_d}{\tau_d^0}$.

C. IoT Power Consumption and Energy Harvesting Model

The energy model of an IoT device relates its communication activity and its states (active, standby, off) to the power consumed. Specifically, in this work, we refer to a class of IoT UEs with an overall power consumption between 1 and 10 mW, such as those whose operations can be sustained by active and passive RF EH [21]. Examples are medical and sensing devices for body area networks, such as [22] and [23]. The model that we adopt is derived from [24] through a measurement-based characterization, and it can be parametrized to account for a wide range of device architectures and types. The expression of the consumed power for an IoT device at x, denoted as $h_{req}(x)$, is

$$h_{req}(x) = e_c\Big(w_1 + w_2 P_{IoT} U_{tx}(x)\Big) \quad (10)$$

with $w_1 + w_2 P_{IoT} U_{tx} \leq 1$. P_{IoT} is the transmit power of the device, which we assume is lower than or equal to a maximum value P_{max}, and e_c is the maximum power consumed by the device. w_1 is the part of the consumed power that does not depend on the load or device configuration, and which models the power consumption due to other tasks (sensing, computing). w_2 modulates the contribution due to communication activity and transmit power. $U_{tx}(x)$ is the

mean fraction of time the device at x is busy transmitting. Note that we assume that downlink processing and sensing are modeled as proportional to the transmitted data, thus via w_2.

Lemma 1. *$U_{tx}(x) = \frac{\bar{\tau}_u}{\tau_u^0}\frac{\phi_u}{f(r,\delta_u)}$ where $f(\cdot,\cdot) \leq 1$ is as in Theorem 1.*

We refer to the Appendix for the proof.

As briefly introduced in the previous section, IoT device can perform both active and passive harvesting. The power harvested via active charging is

$$h_{in}^{active}(x) = \frac{U_d(x) P r^{-\alpha} G \eta \phi_d}{f(r, w_d)} \tag{11}$$

The power harvested via passive charging is given by several contributions:

1) power harvested from transmissions of the serving BS to the other associated users, which depends on the time where the user is not served by the associated BS corresponding to $1 - \frac{1}{f(r,w_d)}$:

$$h_{ass_{BS}}^{passive}(x) = U_d(S(x)) P r^{-\alpha} L \left(1 - \frac{\phi_d}{f(r, w_d)}\right) \tag{12}$$

2) from the transmission of other BS to the cell of the user

$$h_{other_{BS}}^{passive}(x) = I(r(x), \rho)\left(1 - \frac{U_d(S(x))(\phi_d - \eta)}{f(r, w_d)}\right) \tag{13}$$

3) from the transmission of other user in UL in the same cell as the typical user

$$h_{other_{UE}}^{passive}(x) = \left(1 - \frac{U_d(S(x))(\phi_d - \eta)}{f(r, w_d)}\right)\bar{O} \tag{14}$$

where

$$\bar{O} = \frac{(1-\gamma)\delta_u P_{bb} + \phi_u \gamma P_I}{(1-\gamma)\delta_u + \phi_u \gamma} \frac{\lambda_b \pi \alpha}{\alpha - 2} \frac{\bar{\tau}_u}{\tau_u^0}$$

With $h_{in}(x)$ we denote the sum of all of the four quantities for a device located at a distance x from the nearest BS. We define the *energy neutrality* of an IoT device at distance x from the closest BS, denoted as $h_0(x)$, as the difference between the power required to operate and the power harvested. Thus, $h_0(x) = h_{req}(x) - h_{in}(x)$. To obtain which is the percentage of users with a satisfactory energy efficiency, such that $h_0(x)$ is non-negative, we can consider the CDF of harvested energy h_0 as:

$$CDF_h(l) = eval(CDF_r\left(h_0^{-1}(r)\right), l), \quad \text{for all } r \geq 1 \tag{15}$$

where $CDF_r(\cdot)$ is the CDF of the distance between a user and its BS, and $h_0^{-1}(\cdot)$ is the inverse of the function $h_0(x)$, defined previously. This metric is used to evaluate which is the percentage of users with a nonnegative power budget, corresponding to a regime where IoT UEs can harvest all of the consumed power from active and passive RF sources.

To model the BS power consumption, we adopt a flexible and widely-used BS power model from [25], where the total power consumption is expressed as:

$$q_1 + U_d[q_2 + q_3(P - P_{\min})] \tag{16}$$

where q_1 is the static power consumed when the BS is idle, $q_2 U_d$ captures load-dependent power unrelated to transmit power, and $q_3 U_d(P - P_{min})$ models the transmit-power-dependent component, where P is the actual transmit power within $[P_{\min}, P_{\max}]$, and U_d is the downlink utilization. This model accounts for both fixed and traffic-dependent energy consumption and is suitable for analyzing dynamic energy-efficient network strategies.

III. Energy-Optimal Load Shifting

In the following sections, we refer to time as an observation window that is divided into K time slots, where $k \in \{1, ..., K\}$ denotes the label of the k-th slot. Each slot is characterized by the density λ_u^k of installed devices, including both BB and IoT devices. In the no-shifting scenario, we assume that in each time slot, IoT devices transmit the traffic they generate by utilizing the entire available channel capacity. In this case, the per-bit delay corresponds to the value defined in Theorem 1 with $\phi_d^k = \phi_u^k = 1$ for each k. When the active state ϕ_j^k, $j \in \{d, u\}$ of IoTs is not equal to 1, the downlink throughput T_d^k and the uplink throughput T_u^k available to the IoT devices at time slot k, are given by:

$$T_d^k(\phi_d^k) = \frac{\phi_d^k(1-\eta^k)}{\omega_d \tau_d^{0,k}} = \frac{\phi_d^k(1-\eta^k)}{(1-\eta^k)\delta_d \tau_d^{0,k}} = \frac{\phi_d^k}{\delta_d \tau_d^{0,k}}, \tag{17}$$

$$T_u^k(\phi_u^k) = \frac{\phi_u^k}{\delta_u \tau_u^{0,k}}, \tag{18}$$

where $\tau_j^{0,k}$ represents the BB UEs target for the related process and time slot k. Assuming we can over-provision the system when convenient, it is reasonable to assume target different per-bit delays for each time slot. These expressions are based on the fact that, in our system, the per-bit delay is the inverse of the throughput. Additionally, the weights for downlink and uplink scheduling for BB UEs, which are utilized to calculate the per-bit delay, depend on both δ_j and η^k. The traffic sent (resp. received) by IoT devices during time slot k can be expressed as:

$$B_j^k = \gamma^k \lambda_u^k T_j^k(\phi_j^k) \tag{19}$$

To optimize network resource usage and create a more energy-efficient system, our proposed load-shifting mechanism models the ability to redistribute transmission loads across adjacent time slots. In traditional systems, the load generated at time k must be delivered within a certain delay constraint. To capture this, we define a window of size D, which represents the

maximum delay allowed for shifting traffic to neighboring time slots. Considering the periodic nature of a 24-hour day, our model also permits shifting some load from the last time slot to the first. As a result, each time slot can receive load from up to D preceding time slots and forward traffic to up to D subsequent time slots. We define $\phi_{j,0}^k$ as the target ϕ_j^k, which depends on $\tau_j^{0,k}$, and describe the target activation during time slot k, indicating how many bits are generated or received. Denoting by χ_j^k the amount of requested traffic at time slot k:

$$\chi_j^k = \lambda_u^k \gamma^k T_j^k(\phi_{j,0}^k), \tag{20}$$

the total amount of traffic sent (resp. received) in time slot k under the load shifting assumption can be now expressed as:

$$B_j^k = \sum_{i=k-D}^{k-1} B_j^{i,k} + \chi_j^k - \sum_{i=k+1}^{k+D} B_j^{k,i} \tag{21}$$

where $B_j^{i,z}$ with $j \in \{u, d\}$ defines the uplink (resp. downlink) load moved from timeslot i to z. At the end of the observation period consisting of K equal time slots, the total traffic sent (resp. received) will be given by the sum $\sum_{k=1}^{K} B_j^k$.

The proposed modelization implies that some devices may transmit at a throughput lower than their assigned target during specific time slots, while in others they may fully utilize the channel capacity for transmission or reception.

Under the load-shifting assumption, we formulate the energy-optimal problem, which aims to identify for each time slot k, the best system configuration in terms of density of active base stations λ_b^k, splitting factor η^k, BS transmission power P^k, and fraction ϕ_d^k, ϕ_u^k of active IoT devices in downlink and uplink, respectively. Specifically:

Problem 1 (Energy-Optimal Dynamic Network Management)**.**

$$\underset{\substack{\{\lambda_b\},\{P\},\\ \{\phi_u\},\{\phi_d\},\{\eta\}}}{\textit{minimize}} \quad \sum_{k=1}^{K} e_k \lambda_b^k \left[q_1 + \frac{\bar{\tau}_d^k}{\tau_d^0} \left(q_2 + q_3 \big(P^k - P_{\min} \big) \right) \right] \tag{22}$$

Subject to, $\forall t$*:*

$$\begin{aligned}
&\textit{(C1)} \ \ \frac{\bar{\tau}_d^k}{\tau_d^0} \leq 1, \quad \frac{\bar{\tau}_u^k}{\tau_u^0} \leq 1 \\
&\textit{(C2)} \ \ \textit{CDF}_h(0) \leq \iota \\
&\textit{(C3)} \ \ 0 \leq B_d^k \leq (D+1) B_d^k \\
&\textit{(C4)} \ \ \textstyle\sum_{k=1}^{K} \chi_d^k \leq \sum_{k=1}^{K} B_d^k \\
&\textit{(C5)} \ \ 0 \leq B_u^k \leq (D+1) B_u^k \\
&\textit{(C6)} \ \ \textstyle\sum_{k=1}^{K} \chi_u^k \leq \sum_{k=1}^{K} B_u^k \\
&\textit{(C7)} \ \ 0 \leq \phi_u^k, \, \phi_d^k \leq 1 \\
&\textit{(C8)} \ \ 0 \leq \eta^k \leq 1 \\
&\textit{(C9)} \ \ P_{min} \leq P^k \leq P_{max} \\
&\textit{(C10)} \ \ 0 \leq \lambda_b^k \leq \lambda_{b,max}
\end{aligned}$$

where e_k is the energy cost at time slot k. Constraints (C1) ensure that the QoS in downlink and uplink is satisfied for all users, with $\bar{\tau}_d^k$ and $\bar{\tau}_u^k$ as in Equations 5 and 6. Constraint (C2) ensure that the percentage of non-satisfactory harvesting is lower than ι, where $CDF_h(\cdot)$ is given by Equation 15. D represents the maximum allowable delay for the load shift, and the delay constraints are detailed in (C3) and (C5). Constraints (C4), (C6) ensure that the total shiftable traffic in DL and UL is preserved over the observation window, with χ_j^k and B_j^k as in Equations 20 and 21, respectively. Constraints (C7)-(C10) reflect realistic bounds on the values of the active ratios in DL and UL, splitting factor, transmission power, and density of active base stations, respectively. The problem is non-convex and highly non-linear, which makes deterministic approaches impractical. A Genetic Algorithm (GA), was adopted given its effectiveness in related optimization tasks [12]. The computational complexity mainly scales with the number of time slots, since each gene encodes the decision for one slot. To keep the search space manageable, we constrained gene values to the range [0,1] with a discretization step of 0.02, which substantially reduces the number of candidate solutions while preserving solution accuracy.

IV. Numerical Results

In this section we present the optimization results and the optimal load-shifting strategies deducible. We assume that base stations operate in the $1.5\,\text{GHz}$ band and utilize a bandwidth of $50\,\text{MHz}$, in accordance with 5G standards. Unless otherwise specified, the proportion of IoT devices is set to 80% of the total number of UEs, reflecting many current deployment scenarios. Both IoT and BB UEs are assumed to transmit at a power of $0.2\,\text{W}$, and a frequency reuse factor of 3 is adopted. By default, we consider a beamforming gain of 10, constant across the main lobe aperture, which is assumed to be $45°$. The path loss exponent is set to 3, representative of urban environments. The base stations are of the macro type, with transmit power ranging from $1\,\text{W}$ to $11\,\text{W}$. Unless otherwise stated, the target mean per-bit delay in the downlink is set to $10^{-5}\,\text{s}$ for BB UEs and $10^{-3}\,\text{s}$ for IoT UEs. In the uplink, the target delay is $10^{-4}\,\text{s}$ for all UEs, consistent with typical IoT applications such as environmental monitoring [26]. User density is considered to vary between 10^{-2} and 10^{0} users per m^2, modeling scenarios ranging from sparse BB deployments to dense IoT environments. The maximum acceptable proportion of IoT users unable to harvest this minimum energy is limited to 5%. The parameters of the BS energy model are selected to represent the behavior of the high load proportionality (HLP) BS type, with values of $q_1 = 482.3$, $q_2 = 48.23$, and $q_3 = 144.69$. This corresponds to a 75% load proportionality, which can be achieved, for example, through time-domain duty cycling at the subsystem

level [27]. These parameters were specifically chosen to reflect a maximum power consumption of 1500W per BS, which is typical for standalone macro BSs [28]. While, the maximum energy consumed by IoT devices is assumed to be $e_c = 6$mW. The cost of energy per time slot $e_k = 131.47$€/MWh, corresponding to an average daily value for the north of Italy, is assumed to be constant[1]. For $j = d, u$ the target $\phi^k_{j,0}$ is set to 0.08 if $\tau^{0,k}_j = 10^{-6}$, to 0.8 elsewhere. The maximum acceptable delay D for shifting is set to 6. The value of BS density has been varied with a granularity of 10^{-4} m^{-2}, to guarantee good accuracy of the GA search process. We assume the mean number of users per base station to be lower bounded by 5. This models the simple energy-saving strategy common among MNOs, which switch off those BSs that serve very few users to no users at all, as they represent a very high energy cost per user and a small benefit for performance. The initial population size n in the genetic algorithm (GA) was set to 100 chromosomes. This choice struck a good balance between computational load and convergence speed.

A. Analysis

We begin our analysis by examining the system behavior with respect to different target per-bit delays in downlink over a single time slot. Figure 2 shows the optimal energy cost with respect to the overall density of connected UEs. The continuous lines correspond to the novel findings. The green and blue lines correspond to a IoT throughput $T^0_d = \frac{1}{10^{-4}}$. In the green line configuration, the per-bit delay for broadband users is $\tau^0_d = 10^{-6}$, with a delay ratio $\delta_d = 100$. In the blue line configuration, the per-bit delay for BB users increases, while δ_d decreases, resulting in the same effective throughput for IoT devices as in the green configuration.

The red line, on the other hand, maintains the same per-bit delay for BB users as the red configuration but achieves higher throughput for IoT devices, with $\delta_d = 10$. This scenario is particularly relevant to give UEs a greater throughput, which serves as the foundation for load-shifting. In our system, we aim to identify optimal timeslots that can accommodate data originally scheduled for other timeslots. Without adequate available throughput, these timeslots would be unable to handle the additional data. As the required throughput for BB users increases, the system shows a corresponding rise in overall energy consumption compared to the previous scenario. This is due to the need to deploy additional BSs to meet performance requirements. Interestingly, at low user densities, the price increase for improved throughput for BB and IoT UEs is minimal. This indicates that enhancing throughput for users does not significantly impact the network's energy consumption. This is advantageous for MNOs because they

[1]Energy costs can be extracted from the API in https://www.electricitymaps.com/

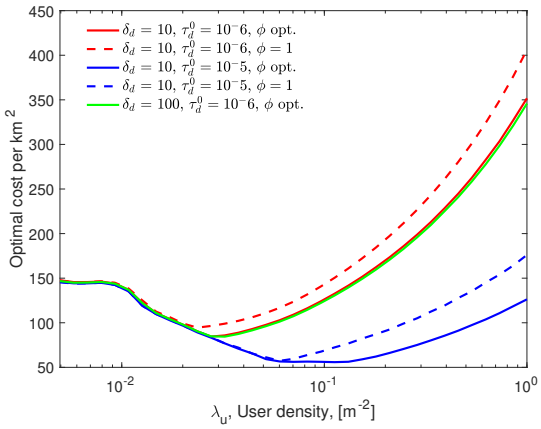

Fig. 2: Energy consumed at the optimum vs. densities of users for different BB and IoT target per-bit delays. $\phi = 1$ refers to the baseline configuration where the IoTs cannot manage their duty cycle tuning ϕ_u and ϕ_d.

can provide users with better service without incurring higher energy costs, thereby preserving energy efficiency. The slight difference in the optimal energy configuration for the two considered values of target per-bit delays is possible thanks to the key role of BS utilization in SWIPT networks. Specifically, at the optimum, in the higher-throughput configuration, the system deploys an additional active BS that enhances the performance of active harvesting because the average distance from the serving BS is decreased. The improved harvesting performances allow for a reduction of η as it is visible in Figure 3. This means that BS can allocate more resources to BB users to satisfy their required QoS. Since the QoS requested is high, they manage to push the downlink utilization to 1. This feature is fundamental to maximizing the passive and active harvesting performances. In fact, from its value depend both the active and the passive harvesting: a higher utilization allows for a more efficient harvesting, as it is evident from Equations 11, 12, 13, and 14. In the lower-throughput configuration, it is not convenient to activate more BS since it will translate into an average lower BS downlink utilization, with negative consequences on the harvesting performances. This explains why the DL utilization in both configurations is always at the maximum (even if the energetic model is high-load proportional). Since the amount of bits received by IoTs is constant across all configurations, the red and green lines coincide because IoTs do not require an additional channel for their transmissions.

The dashed lines correspond to the baseline scenario where IoT devices are always active both in downlink and uplink (i.e., $\phi_u = \phi_d = 1$). When IoT devices cannot tune their duty cycle, the energy consumed by the networks increases for higher densities because the BSs are underused even if

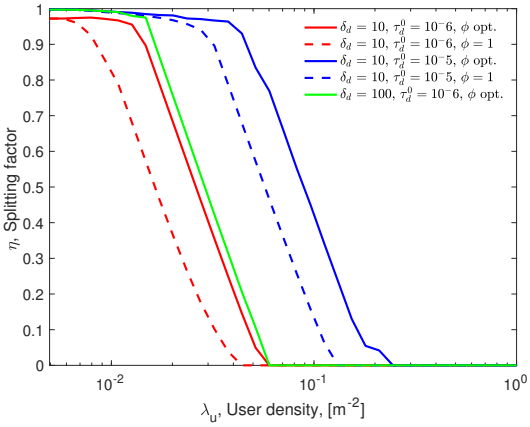

Fig. 3: Splitting factor at the optimum vs. densities of users.

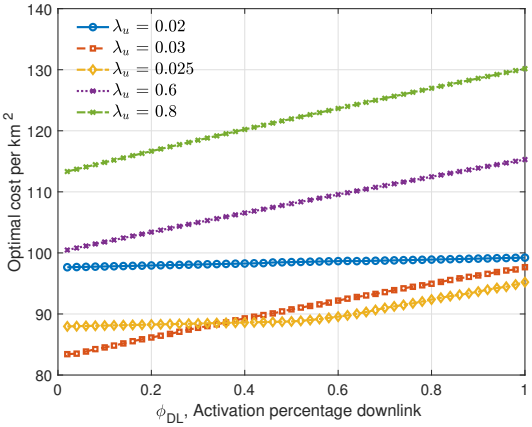

Fig. 4: Energy consumed at the optimum vs. ϕ_d for different users' configurations.

the passive harvesting is more effective due to the longer duration of transmissions (as it is evident from the earlier decrease in η visible in Figure 3). On the contrary, the dynamic increase of the active users achieves a balance point where the power harvested from surrounding users equals the power required to operate the device, without requiring additional energy. This allows the system to maintain the splitting factor near zero for a wider range of densities, as devices can adapt their consumption and reduce overall energy consumption. Such a behavior is only possible through flexible control over device activation. Furthermore, the divergent behavior of the curves for $\lambda_u > 3 \times 10^{-2}$, is clarified by Figure 4, which shows the energy cost with respect to the downlink activation of IoTs, for different densities of installed users. For high densities, the activation time impacts the network energy consumption. Indeed, reducing the load in that specific density (i.e., time slots) could allow for saving up to 15% of the energy cost. The load can then be shifted during a less crowded time slot in which the impact of the load is lower. To explore this possibility, we focused on an average daily traffic profile derived from [29]. The system's behavior is illustrated in Figure 5. In this figure, the red color indicates the time slots where over-provisioning the network is advantageous, as shown in the previous analysis. Consistent with the trend observed in Figure 4, the most congested time slots correspond to the highest gains in energy costs (i.e., the amount of load shifted). Reducing the duty cycle during those time slots allows for higher relative savings. Consequently, the load is redirected to lighter slots, which are now more heavily utilized. This is possible thanks to network over-provisioning.

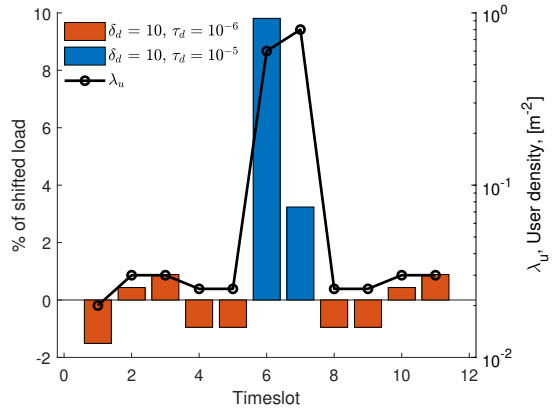

Fig. 5: Optimal shifting strategy for a 24-hour observation window with a time slot duration of 2 hours. The black line represents the considered traffic profile.

V. Conclusions

This paper presents an analytical framework to optimize energy-efficient load shifting in SWIPT-enabled networks, balancing operational costs, network performance, and device lifetime. Our results demonstrate that strategic over-provisioning during low traffic periods can improve QoS without increasing energy consumption. We highlight the nonlinear and often counterintuitive relationship between traffic demand and energy use, enabling novel energy-saving strategies. The proposed approach empowers MNOs to dynamically redistribute load to intervals in which the energy consumed is less sensitive to traffic demand fluctuations, while maintaining user-perceived QoS. This is particularly impactful for IoT devices capable of adapting their communication schedules. It is important to note that savings vary based on the traffic profile considered, meaning different scenarios can result in different savings. A promising direction for future work is the integration of real-time green energy availability and storage conditions into the load shifting framework, which could further enhance the sustainability and operational efficiency of SWIPT networks.

ACKNOWLEDGMENTS

The paper was supported by the UNITY6G project funded by the European Union's Horizon Europe Research and Innovation Programme under Grant N.101192650 and the DFG within the DyMoNet project.

REFERENCES

[1] R. Singh *et al.*, "Internet of things (iot): Opportunities, issues and challenges towards a smart and sustainable future," *Journal of Cleaner Production*, vol. 274, p. 122878, 2020.

[2] O. L. López *et al.*, "Energy-sustainable iot connectivity: Vision, technological advances, and challenges," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 2609–2651, 2023.

[3] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5g: physical and mac-layer solutions," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 59–65, 2016.

[4] R. Kamran, S. Kiran, P. Jha, A. Karandikar'T, and P. Chaporkar, "Green 6g: Energy awareness in design," in *2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS)*, 2024, pp. 1122–1125.

[5] Z. Chang, Z. Han, and W. Saad, "Energy harvesting and power management for iot devices in the 5g era," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 91–97, 2021.

[6] G. M. Razavi and A. H. Nejati Sharif, "Intelligent networking for energy harvesting powered iot systems," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 7, no. 1. ACM, 2023, pp. 1–28.

[7] K. Nguyen and M. Ghovanloo, "Making sensor networks immortal: An energy-renewal approach with wireless power transfer," *IEEE/ACM Transactions on Networking*, vol. 20, no. 6, pp. 1748–1761, 2012.

[8] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 902–912, 2014.

[9] Y. Zhang *et al.*, "Energy efficiency optimization for swipt-enabled iot network with energy cooperation," *IEEE Access*, vol. 10, pp. 123 456–123 468, 2022.

[10] R. Ma, J. Tang, X. Y. Zhang, W. Feng, D. K. C. So, C.-B. Chae, K.-K. Wong, and J. A. Chambers, "Simultaneous wireless information and power transfer in iot-based scenarios: Architectures, challenges, and prototype validation," *IEEE Wireless Communications*, vol. 30, no. 1, pp. 36–43, 2023.

[11] T. A. Khaleel, F. A. Mustafa, and M. F. Khattab, "Applications of sensor networks and remote sensing in environmental sustainability: A review," in *2022 International Conference on Engineering & MIS (ICEMIS)*, 2022, pp. 1–3.

[12] B. Boi *et al.*, "Green operations of swipt networks: The role of end-user devices," *arXiv preprint arXiv:2312.08232*, 2024, available at: https://arxiv.org/abs/2312.08232.

[13] G. Rizzo *et al.*, "Energy-optimal ran configurations for swipt iot," in *2022 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2022, pp. 1–6. [Online]. Available: https://arodes.hes-so.ch/record/11333/files/Rizzo_2022_energy-optimal_RAN.pdf

[14] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for swipt systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2082–2085, 2015.

[15] J. Wu, E. W. M. Wong, Y.-C. Chan, and M. Zukerman, "Power consumption and gos tradeoff in cellular mobile networks with base station sleeping and related performance studies," *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 4, pp. 1024–1036, 2020.

[16] Y. Wu, Y. He, L. Qian, J. Huang, and X. Shen, "Joint scheduling and power allocations for traffic offloading via dual-connectivity," *arXiv preprint arXiv:1509.09241*, 2015.

[17] P. Sroka *et al.*, "Policy-based traffic steering and load balancing in o-ran-based vehicle-to-network communications," *IEEE Access*, pp. 9357–9376, 2024.

[18] L. Kundu *et al.*, "Towards energy efficient ran: From industry standards to implementation practice," *arXiv preprint arXiv:2402.11993*, 2024.

[19] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3622–3636, jul 2013. [Online]. Available: https://doi.org/10.1109%2Ftwc.2013.062413.122042

[20] G. Rizzo, M. A. Marsan, C. Esposito, and B. Boi, "Green operations of swipt networks: The role of end-user devices," *IEEE Transactions on Green Communications and Networking*, pp. 1–1, 2025.

[21] P. N. Alevizos and A. Bletsas, "Sensitive and nonlinear far-field rf energy harvesting in wireless communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3670–3685, 2018.

[22] P. K. Miriyala, P. N. Srinivas *et al.*, "On-chip 5&6-ghz rf energy harvesting system for implantable medical devices," in *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024, pp. 1–5.

[23] N. U. Khan, F. U. Khan, M. Farina, and A. Merla, "Rf energy harvesters for wireless sensors, state of the art, future prospects and challenges: a review," *Physical and Engineering Sciences in Medicine*, vol. 47, no. 2, pp. 385–401, 2024.

[24] Özen Özkaya and B. Örs, "Model-based, fully simulated, system-level power consumption estimation of iot devices," *Microprocessors and Microsystems*, vol. 105, p. 105009, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141933124000048

[25] M. S. Mushtaq, S. Fowler, and A. Mellouk, "Power saving model for mobile device and virtual base station in the 5G era," in *IEEE ICC*, Paris, France, May 2017, pp. 1–6.

[26] W. Lu, X. Xu, G. Huang, B. Li, Y. Wu, N. Zhao, and F. R. Yu, "Energy Efficiency Optimization in SWIPT Enabled WSNs for Smart Agriculture," *IEEE Trans. Ind. Inf.*, vol. 17, no. 6, pp. 4335–4344, Jun. 2021.

[27] J. Lin, Y. Chen, H. Zheng, M. Ding, P. Cheng, and L. Hanzo, "A data-driven base station sleeping strategy based on traffic prediction," *IEEE Transactions on Network Science and Engineering*, 2021.

[28] X. Ge, J. Yang, H. Gharavi, and Y. Sun, "Energy efficiency challenges of 5g small cell networks," *IEEE communications Magazine*, vol. 55, no. 5, pp. 184–191, 2017.

[29] M. Meo, D. Renga, and Z. Umar, "Advanced sleep modes to comply with delay constraints in energy efficient 5G networks," in *VTC Spring*, 2021, pp. 1–7.

APPENDIX

Proof sketch of Lemma 1. To derive the uplink utilization $U_{tx}(x)$ of IoT located at distance x from the associated BS, we start by considering the time the device is busy transmitting, which corresponds to the ratio ϕ_u over the entire uplink time. Similar to base stations, the energy consumption of an IoT device is a rescaling of the maximum energy consumed when the IoT device is transmitting at its peak capacity. The rescaling factor is determined by a fixed quantity that represents the energy consumed when the IoT is not transmitting data and a variable component that relies on both the transmission power and the IoT's utilization. The latter can be derived from the mean uplink utilization of BS $\frac{\bar{\tau}_u}{\tau_u^0}$. Given that BS resources are distributed among $f(r, \delta_u)$ users and the uplink weight of IoTs is 1, we can calculate the amount of resources (time) allocated by the BS to IoTs, which corresponds to the duration in which the device is actively transmitting information. The actual transmitting time is then multiplied by the fraction of time the IoT scheduled to use for transmitting, i.e., ϕ_u □