Energy Efficient Green Wireless Communication Systems with Imperfect CSI and Data Outage

Charilaos C. Zarakovitis, Qiang Ni and John Spiliotis

Abstract—**Modern applications involve green communication technologies motivating well optimisation in the power–limited regime. In comparison to most of existing related work that assumes perfect *channel state information* (CSI) is always available, which is unfortunately not true in reality, this work focuses on an optimal energy efficient solution for resource allocation in multiuser *orthogonal frequency division multiple access* (OFDMA) networks in the presence of imperfect CSI and data outage conditions. Particularly, in view that wireless channel conditions, circuit power consumptions and users' *quality–of–service* (QoS) requirements are heterogeneous in nature, we enable attractive tuning options by letting energy efficiency optimisation objective to assign weights to each allocation link. Also, we interpret effects of data outage due to imperfect CSI using a profound insight on the monotonicity of noncentral chi-squared inverse distribution function, which reveals that our design complies with expected physics and mechanics of conventional energy efficiency approach and that it can be successfully degenerated to the energy efficiency model with perfect CSI. Furthermore, we formulate a mixed combinatorial problem towards maximising the energy efficiency subject to a minimum QoS requirement, channel interference and transmitting power constraints. The problem is transformed into an equivalent quasiconcave problem with respect to power, and concave problem with respect to the subcarrier indexing coefficients using the concept of subcarrier time–sharing. We optimise through a simple and versatile methodology, which uses standard–Lagrangian optimisation technique to obtain joint dynamic subcarrier and adaptive power allocations by means of final formulas. We also examine key properties of the introduced optimal solution in terms of implementation convergence and complexity, level of optimality, and impact of imperfect CSI coefficients and circuit power on network performance. The simulation results demonstrate the effectiveness of our allocation scheme for achieving higher energy efficiency performance with the guaranteed QoS support and lower complexity than existing approaches especially when perfect CSI is not available.**

Index terms— Convex optimisation, CSI, data outage, energy efficiency, multiuser, multicarrier, OFDMA, QoS, resource scheduling

I. INTRODUCTION

The rapid growth of requirements for high–speed wireless communication services and ubiquitous access escalates dramatically energy consumption. In addition to this issue, slow

C. C. Zarakovitis and Qiang Ni are with the School of Computing & Communications, Lancaster University, LA1 4WA, UK (e-mail: {c.zarakovitis, q.ni}@lancaster.ac.uk).

John Spiliotis is with the School of Applied Mathematics & Physical Sciences, National Technical University of Athens (NTUA), Iroon Polytexniou 9, 15780, Zografou, Athens, Greece (jspil@math.ntua.gr).

advancements on energy–storage technology bound the power supply of portable devices with limited battery capacity [1]. Such matters convey critical financial and environmental concerns, like high operation expenditure and large amount of greenhouse gas emissions [2]. A promising solution capable to guarantee certain levels of energy efficiency and *quality–of–service* (QoS) support is the intelligent management of wireless resources considering features of green networking and *orthogonal frequency division multiple access* (OFDMA) technologies [3].

A. Related Work and Open Challenges

Traditionally, OFDMA resource schedulers attain the required QoS by either maximising sum–throughput or minimising total transmission power, e.g., [4]–[5]. Even so, [6] reveals that such efforts are far apart the optimality from the perspective of energy efficiency defined as the ratio between offered throughput over transmit and circuit power (in *bit/Joule*). Therefore, [7]–[13] strove to optimise energy efficiency using power adaptations within multi–user channel coordination to improve QoS maintenance. Within same direction, [14]–[17] examine the tradeoff between spectral and energy efficiency to reveal fundamental impacts of channel power gain and circuit power on energy performance. Further insights are given in [18]–[22], which show that the amount of data bits successfully delivered to the receiver for each energy unit used for transmission should be maximised on joint subcarrier and transmitting power level considerations. The fundamental assumption of all above studies is that *channel state information* (CSI) is always perfectly known. In practice this is challenging because schedulers cannot perfectly obtain the CSI due to noisy channel conditions, delayed CSI feedback and/or channel outage, etc. With imperfect CSI, data could be corrupted if scheduled data rate exceeds the maximum channel capacity resulting in poor energy efficiency. For this reason, recent studies in [23]–[26] account properties of Marcum Q- and Bessel functions to interpret imperfect CSI in energy efficiency. Particularly, [23] and [24] consider quantisation errors and/or limited feedback information to improve energy efficiency in coordinated multi-cell networks and cognitive radio systems, respectively. Deterministic bounded channel modelling is adopted in [25] and [26] to address worst case beamforming strategies that optimise downlink energy efficient transmissions in single-cell and coordinated multi-cell *multiple-inputmultiple-output* (MIMO) networks, correspondingly. It is demonstrated that imperfect–CSI–awareness can bring energy improvements compared to previous attempts with perfect CSI considerations. An issue however, is that

imperfect CSI coefficients in [23]–[26] are interpreted using exponential–type approximations, which can be calculated numerically either with infinite series expansions or through recursive relations. Such approximations very often require a large number of iterations to yield sufficient accuracy for very small error probabilities, whereas asymptotic expressions may give values that lie above or below the exact value of the function. For example, the generalised Marcum Q-function in [23]–[26] is expressed as an integral with finite limits, which cannot be solved analytically. Hence, complex numerical integration techniques have to be applied introducing thereby, approximation errors over the large domain. How to avoid such errors and improve recursive procedures towards more optimal solution with the existence of imperfect CSI is a non–trivial question and remains challenging.

Furthermore, since the energy efficiency optimisation problems lead to fractional models, relevant literatures [7]–[26] apply *fractional programming* to facilitate dual Lagrangian optimisation, which exploits the underlying linearity of the linear fractional model. The resulting algorithms can be seen as specialisations to Dinkelbach-type algorithm for generalised fractional programs [27], which parametrises the master problem into a series of concave secondary problems that are solved using dual decomposition, where each secondary problem is further decomposed into a number of similar subproblems. The challenge is that dual methods often lead to sub–optimal solutions due to uncontrollable duality gap [28]. Given that the gap cannot be explicitly calculated [28], dual methods aim to reduce the gap assuming that each concave secondary problem has a strictly feasible point such that the Slater's condition is satisfied (see [18], [23], [29] for example). Also, dual methods call for an application of rather complex searching processes for obtaining the subgradients of each possible optimal point (e.g., ellipsoid method). Such processes increase complexity and cannot always guarantee convergence since they may result in more than one subgradient or even non [28]. Commenting on the high complexity and potential inaccuracy caused by the duality gap from those Lagrangian dual methods, a recent study in [30] attempts to use Lagrangian optimisation with numerical approach. It shows that the *Karush-Kuhn-Tucker* (KKT) optimality conditions of the energy efficiency problem lead to transcendental equations in the form of $(1 + x) \cdot \ln (1 + x) + x = 0$, which are challenging to be solved analytically. Thus, authors propose a numerical method, which achieves higher performance than existing dual approaches by means of convergence, complexity and accuracy. Despite the promising results in [30], how to exploit assets of standard–Lagrangian optimisation analysis towards further energy efficiency improvements is still in early stage.

B. Motivation and Contributions

This study identifies the importance of recent attempts in [23]–[26] and [30] to step forward by investigating new potentials for improving energy efficiency through standard–Lagrangian analysis in OFDMA systems with the existence of imperfect CSI. Our viewpoint is that new analytical solutions, which get rid of the duality–gap and subgradient seeking processes can improve optimality, convergence and complexity

of the energy efficiency problem. In addition, the negative impact of imperfect CSI can be addressed through poring into fundamental properties of noncentral chi-squared inverse *cumulative distribution function* (c.d.f.), which can address the issues caused by Marcum Q- and Bessel functions. To this end, we focus on investigating potentials of an uncommon figure of merit, known as *weighted sum energy efficiency* (WSEE), which allows to control the energy efficiency of each allocation link over the available spectrum. Compared to conventional *average sum energy efficiency* (ASEE), which considers the energy efficiency of the entire network, WSEE is useful in heterogeneous networks wherein different users have diverse QoS requirements and channels have different quality levels. The key contributions of this paper are summarised as follows:

- In view that channel conditions, circuit power consumptions and users' QoS vary in nature, we focus on maximising WSEE to enable attractive tuning options in energy efficient communication systems. Related efforts mainly investigate ASEE performance [7]–[15], [18]–[26], while the works that consider the scenario similar to ours are [16]–[17], to the best of our knowledge. However, [16]–[17] optimise using power adaptation only, while subcarrier assignments are assumed to be fixed. Instead, this work proposes a joint subcarrier and power allocation solution. The joint approach captures different features of energy performance than unilateral power optimisations. Furthermore, we design WSEE by accommodating the case when CSI is not perfectly available, which had not been considered previously.

- We investigate the effect of data outage due to imperfect CSI on energy efficient allocations. This topic is still in its infancy as most related work (e.g. [7]–[22], [30]) assumes that CSI is perfectly available, which is unrealistic and may lead to poor performance. Although some recent works [23]–[26] consider studying the impact of imperfect CSI on the ASEE metric, they use recursive methods to compute the wireless channel, which introduce statistical approximation errors. We contribute by expressing for the first time the statistics of imperfect CSI in WSEE using inverse chi-squared c.d.f., which avoids recursive computations and can improve approximation accuracy compared to [23]–[26]. This is achieved through a profound insight on the monotonicity of noncentral chi-squared inverse c.d.f. with respect to its noncentral parameter. To the best of our knowledge, such insight has not been examined previously. Instead, the monotonicity has been examined with respect to the random variable (not with respect to its noncentral parameter), e.g., [31]–[32]. We expect our insight to be a useful tool for the green communication and networking research field as well as wider communication engineering fields.

- Different from dual Lagrangian approach [7]–[26], we apply an intelligent solution methodology to enable standard–Lagrangian optimisation. The idea relies on Maclaurin theorem, which incurs the polynomial form of transcendental equations issued by the fractional origin of energy efficiency objectives. We show that truncation error of our optimal solution can be expressed in

closed–form, while duality gap in related work is non-tractable and may impact optimality. Moreover, it is the first attempt to consider that subcarriers can be time–shared in energy efficient scheduling. Although, time–sharing is a well-studied technique capable to reduce implementation complexity in spectral efficiency problem, e.g., [4]–[5], [31]–[32], it has been rarely used in the energy efficiency area, whereas conventional approach fosters to seek for the subgradient of each dual optimal subcarrier allocation, i.e., [7]–[26]. We also contribute by exploring the boundary points of our optimal solution and its key properties with proofs regarding its monotonicity, complexity, feasibility, implementation and optimality.

Simulation comparisons demonstrate that the proposed resource scheduling approach has significantly lower complexity than dual approaches and achieves higher optimality. It is also shown that the introduced modelling of energy efficiency with imperfect CSI can be successfully degenerated to the energy efficiency model when perfect CSI is available. The new objective and solution methodology can be applied to a large class of energy efficiency problems under various system constraints, and offer useful insights to theoretic and algorithmic developments for wireless utility maximisation.

Remark: The main target of this work is to investigate the potential gain on energy efficiency from exploiting the statistical CSI when CSI is not perfectly known. This is achieved by the design and the solution of an energy efficient allocation problem subject to outage constraints. By contrast, our previous work [33] relied on a different energy efficiency metric (i.e. ASEE) than the energy efficiency we address in this paper (i.e. WSEE). As discussed previously the two definitions have fundamental structure differences. Thus, the corresponding properties/proofs presented in this work and the work in [33] are radically different. Apart from the different energy efficiency definitions and properties, the key differences between the current work and [33] lie in the optimisation problem, the solution approach, the algorithm design, and the performance analysis. In this paper, we present the properties not only of the energy efficiency objective but also of the joint solution. Along with that, we present the implementation process of our proposal as well as extended performance comparisons with relevant studies through simulations. It is also noted that our previous work [32] addresses the impact of imperfect CSI relying on conventional Marcum Q-function properties. Instead, for the first time in this paper we investigate the energy efficiency optimisation with the existence of imperfect CSI using attributes on the monotonicity of noncentral chi-squared inverse c.d.f.

The remainder of this paper is structured as follows. *Section II* presents the OFDMA system model, with *Section III* to introduce the new energy efficiency objective with key properties. *Section IV* formulates the optimisation problem with *Section V* to present the joint optimal allocation solution. *Section VI* provides main attributes of the joint solution, with *Section VII* to present various simulation comparisons with relevant schemes. Finally, *Section VIII* concludes the paper.
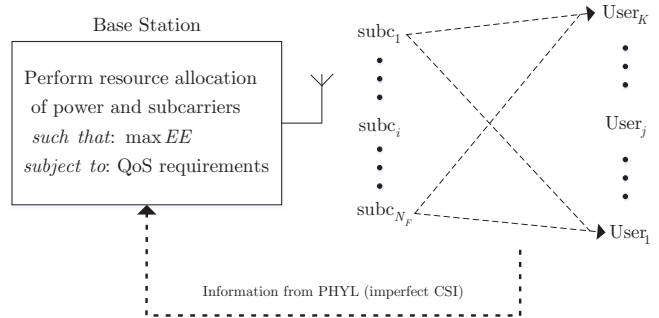
Fig. 1: Multiuser OFDMA system model.

II. SYSTEM MODELLING

We focus on downlink transmissions of a single-cell OFDMA system shown in *Fig. 1*. In the system, K total number of mobile users (each with a single antenna) share N_F total number of orthogonal subcarriers of total bandwidth BW. Users feedback CSI to a *base station* (BS), which is equipped with a single antenna and determines the subcarrier and power allocation decision. The CSI is imperfect due to channel feedback delays, transmission errors, etc., and it is estimated at the BS prior to the allocation decision of each transmission circle. The allocation decision aims to maximise energy efficiency subject to the minimum QoS requirements of each user, interference constraints among N_F subcarriers and the transmitting power limit.

A. Wireless Channel Modelling & Estimation Process

The wireless channel is frequency–selective with *complex circularly symmetric Gaussian* (CCSG) noise of spectral density N_0. The channel estimation is performed in *time division duplex* (TDD) operation through uplink dedicated training blocks sent by all the K users to the BS. Let h_{ij} to denote the complex Gaussian and *identically independent distributed* (i.i.d) actual channel gain of user $j = 1,...,K$ on subcarrier $i = 1,...,N_F$. Then the received OFDM symbol can be represented by

$$y_{ij} = h_{ij} \cdot x_{ij} + z_{ij}. \tag{1}$$

Variable x_{ij} in (1) signifies the transmitted *orthogonal frequency division multiplexing* (OFDM) symbol to user j on subcarrier i, while $z_{ij} \sim \mathcal{CN}\left(0,\sigma_Z^2\right)$ symbolises the corresponding CCSG channel noise with zero mean and variance $\sigma_Z^2 = \frac{BW \cdot N_0}{N_F}$. Also, let the estimated channel gain coefficient $\hat{h}_{ij} = h_{ij} + \Delta h_{ij}$ to represent the imperfect CSI, with Δh_{ij} the i.i.d channel gain estimation error term and $\sigma_{\Delta h}^2$ the channel gain error variance. Then $h_{ij} \sim \mathcal{CN}\left(\hat{h}_{ij}, \sigma_{\Delta h}^2\right)$ and the evaluation of actual channel gains $h = \{h_{ij}\}$ can be performed according to the estimation process proposed in [31]–[32], which ensures that:

1) channel gain estimation error terms $\Delta h = \{\Delta h_{ij}\}$ are correlated on different subcarriers with same channel gain error variance $\sigma^2_{\Delta h}$ and

2) Δh and estimated channel gain realisations $\hat{h} = \{\hat{h}_{ij}\}$s are uncorrelated with zero covariance.

B. Physical Layer Modelling

In view of (1), we define the system transmitting power allocation policy by $\mathcal{P}\left[\hat{H}_{N_F \times K}\right] = [p_{ij}]$, with matrix element the averaged (or expected) instantaneous transmitting power $p_{ij} = E\left[|x_{ij}|^2\right], p_{ij} \geq 0, \forall i, j$ [1] and $\hat{H}_{N_F \times K} = \left[\hat{h}_{ij}\right]$ the imperfect CSI knowledge matrix from physical layer. In addition, we represent the subcarrier allocation policy by $\mathcal{S}\left[\hat{H}_{N_F \times K}\right] = [s_{ij}]$, with $s_{ij} \in \{0, 1\}$ the subcarrier index, i.e., $s_{ij} = 1$ when subcarrier i is allocated to user j and $s_{ij} = 0$, otherwise. In our system the instantaneous transmitting power p_{ij} can be adjusted according to the combined subcarrier gains $\{\hat{h}_{ij}\}$ by using *M-ary Quadrature Amplitude Modulation* (M-QAM) [34]. Thus, the maximum instantaneous channel capacity c_{ij} can be then computed as $c_{ij} = BW \cdot \log_2\left(1 + \left(\eta \cdot p_{ij} \cdot |h_{ij}|^2\right)\right)$, with variable $\eta = -\frac{1.5}{\sigma^2_Z \cdot \ln(5 \cdot BER)}$ to be added for notational brevity, $BER = 0.5 \cdot erf\left(\frac{E_b}{N_0}\right)$ to define the normalised *bit-error-ratio* (BER), E_b the energy-per-bit and $CNR = \frac{p_{ij} \cdot |h_{ij}|}{\sigma^2_Z}$ the *channel-to-noise-ratio* (CNR). Furthermore, we represent the throughput allocation policy by $\mathcal{R}\left[\hat{H}_{N_F \times K}\right] = [r_{ij}]$ with matrix element the instantaneous scheduled throughput $r_{ij} \leq c_{ij}, \forall i, j$, which is a function of actual channel gains $h = \{h_{ij}\}$ that are unknown to the BS. Having specified power, subcarrier and throughput policies we define the instantaneous energy efficiency as the ratio of achievable instantaneous scheduled throughput over the corresponding power consumption, i.e., $EE_{ij} = s_{ij} \cdot r_{ij}/\left(\zeta \cdot s_{ij} \cdot p_{ij} + P^C_{ij}\right)$. Particularly, the power consumption model of the BS is formulated in linear fashion[2] $\zeta \cdot s_{ij} \cdot p_{ij} + P^C_{ij}$ [7]–[18], where coefficient ζ refers to the power consumption due to amplifier inefficiency, feeder losses, etc., and P^C_{ij} symbolises the digital power consumed at the BS's transmitter and the distant receiver for each allocation link due to signal processing, battery backup, etc. Hence, EE_{ij} can be quantified as the number of successfully delivered data bits to user j on subcarrier i per unit power (in *bit per Joule*) and it is equivalent to the throughput weighted by the inverse of the sum of circuit and transmitting power. Furthermore, the scheduling process is constrained on $E\left[\sum_{i=1}^{N_F}(s_{ij} \cdot r_{ij})\right] \geq q_j, \forall j$ to guarantee that each user j

meets its minimum QoS rate requirement q_j. Also, scheduler ensures the total transmitting power upper bound at BS is met, i.e., $E\left[\sum_{j=1}^{K}\sum_{i=1}^{N_F}(s_{ij} \cdot p_{ij})\right] \leq P_{TOTAL}$, where P_{TOTAL} denotes the total available power supplied at the BS. Finally, we restrict channel interference by considering $\sum_{j=1}^{K} s_{ij} = 1$, $\forall i$, which imposes that each subcarrier can be only assigned to one user at a time.

III. DEFINITION AND PROPERTIES OF EFFECTIVE–WEIGHTED ENERGY EFFICIENCY PERFORMANCE METRIC

This Section designs an attractive energy performance metric to account practical effects of data outage and imperfect CSI. The performance metric considers weights to facilitate addressing energy efficiency achieved in each allocation. Key structural determinants of the proposed design are also clarified to offer a unique insight on the monotonicity of the noncentral chi-squared inverse cdf.

A. Definition of Weighted Sum Energy Efficiency

Most literature defines energy efficiency as the ratio between network sum rate and corresponding power consumption, i.e., [7]–[26],

$$\text{ASEE} = \frac{\sum_{j=1}^{K}\sum_{i=1}^{N_F} r_{ij}}{\sum_{j=1}^{K}\sum_{i=1}^{N_F}(\zeta \cdot s_{ij} \cdot p_{ij}) + P^C}, \quad (2)$$

with $P^C = \sum_{j=1}^{K}\sum_{i=1}^{N_F} P^C_{ij}$ the overall circuit power consumption at the BS. However, ASEE in (2) is not able to manage the individual energy efficiencies of each (i, j) allocation link, which are needed when different QoS requirements, power consumption levels and spectrum usage are demanded. To solve this issue, we adopt the weighted sum of the individual energy efficiencies as:

$$\begin{aligned}\text{WSEE} &= \sum_{j=1}^{K}\sum_{i=1}^{N_F}(w_{ij} \cdot EE_{ij}) \\ &= \sum_{j=1}^{K}\sum_{i=1}^{N_F}\left(w_{ij} \cdot \frac{r_{ij}}{\zeta \cdot s_{ij} \cdot p_{ij} + P^C_{ij}}\right),\end{aligned} \quad (3)$$

with $w_{ij} \geq 0$ the weight of each (i, j) link. WSEE provides more degrees of freedom for system design than conventional ASEE because the corresponding weights in (3) can give priority to specific (i, j) allocation links [14], [37]–[39]. For instance, user weighting has been extensively applied in networks with heterogeneous QoS applications to prioritise the most energy–hungry users (with higher QoS than others), which permits obtaining resource allocation solutions that privilege the energy efficiencies for such users. In addition, subcarrier weighting has been accounted in networks, where subcarriers are reserved for different pricing policies and/or spectrum usage [14], [37]. In such networks, subcarrier–dependent weights are used to opportunistically allocate more power to subcarriers (or groups of subcarriers) experiencing a good channel state, or alternatively, to force a uniform power split in order to avoid well-known phenomena with negative effects on network's performance such as sidelobes in OFDM, high value of *peak-to-average-power-ratio* (PAPR), high *out-of-band-power* (OBP), etc. [40]–[41]. Remark that in reality the reliability of subcarriers depends not only on their physical channel coefficients (which can be seen as physical weights) but it is also affected by the signal processing process and

---

[1] The expectation operator $E.[\cdot]$ refers to the average of a quantity over the ergodic realisations, e.g., $\{x_{ij}\}$, $\{h_{ij}\}$, $\{\hat{h}_{ij}\}$, etc. To avoid confusion, from this point forward we omit its characterisation as "averaged" or "expected". Also, notation $|\cdot|^2$ indicates the absolute squared value of a complex number, e.g., $|x_{ij}|^2 = x_{ij}[x_{ij}]^\dagger$, and $[\cdot]^\dagger$ is the complex conjugate [28].

[2] In practical setting, power consumption can be modelled in more sophisticated fashion, where P^C_{ij} can vary with throughput and ζ can change with the bandwidth used, e.g., [35]–[36]. Note that more digital processing is done at the receiver than at the transmitter, whereas analogue circuit power consumption is mostly accounted by coefficient ζ. How to model more complex power model is beyond the scope of this study, which aims to pore into the impact of imperfect CSI on energy efficiency. For the ease of presentation and the fair comparisons with conventional schedulers we adopt the same type of power dissipation modelling as our related work (see [7]–[18]).

circuit of the receiver. For example, during the analogue-to-digital conversion of the OFDM signal, quantisation noise is introduced by the modulators, e.g., QAM in our case. The noise degrades signal reliability by means of reducing the CNR and/or increasing the BER in some subcarriers [42], which can give rise to high sidelobe, PAPR, OBP, etc. To improve reliability the likelihood of the received symbols on subcarriers for decoding can be adjusted by weighting subcarriers based on real-valued coefficients (e.g. CNR, BER), which can be calculated[3] from the output noise spectrum of the modulator using, e.g., soft-decision Viterbi process [43]. As a result, the weights in WSEE can be adjusted considering multiple criteria that lead to more realistic interpretations of each subcarrier's quality and hierarchy than considering the physical channel coefficients only. This can be particularly useful in 5G systems, where different groups of frequency bands are required to rely on separate *radio frequency* (RF) chains, e.g., cellular *multiple-input multiple-output* (MIMO) *cognitive radios* (CRs) utilising carrier aggregation and millimeter waves. In this situation, the RF chains may have different energy consumption levels, and a proper adjustment of the weights permits obtaining different system tradeoffs to privilege the most energy–consuming sub-carriers in selected users [37], [44]. Last, but not least, proper adjustments of joint user and subcarrier weights in WSEE can control the spread of the individual energy efficiencies over the available spectrum. Thereby, energy concentration or energy dispersal (entropy) can be controlled over the available spectrum. As shown later in our simulation results, the option of controlling the entropy energy of wireless transmissions improves performance, which matches with well-known laws of physics.

Hence, while ASEE in (2) accounts for the energy efficiency of the entire network, WSEE in (3) provides additional tuning options to optimise the energy efficiencies of each (i, j) link by considering the heterogeneous user requirements and subcarrier fidelity. Tuning options in WSEE can promote employing dynamic power dissipation, which is not available in ASEE. For example, P_{ij}^C in WSEE can vary with throughput, which is particularly convenient upon considering heterogeneous power consumptions due to different hardware equipment of users. Also, weights in WSEE can facilitate applying game theoretical concepts such as Nash bargaining [45], Max-Min fairness, etc. to enable the fair cooperation among users, which is more effortful in ASEE.

This work focuses on adopting the energy efficiency metric WSEE as defined in (3). Notice that both ASEE in (2) and WSEE in (3) have been mostly formulated assuming perfect CSI is always known, i.e., in the fashion of $EE = \frac{\text{throughput}}{(\text{radiated + circuit) power}}$ [7]–[22], [27], [30]–[33]. However, under imperfect CSI, the actual channel gain realisations $h = \{h_{ij}\}$ in (1) are unknown to the BS and therefore, energy efficiency is random. To address this issue the next sub–Section designs energy efficiency in the fashion of $\overline{EE} = \frac{\text{effective–throughput}}{(\text{radiated + circuit) power}}$, to represent the *effective* WSEE considering the impact of imperfect CSI.

---

[3]Discussing actual policies to calculate user and subcarrier weights is outside the scope of this work. For example, see [37]–[39], [40]–[41] and references within.

B. Design of Effective–Weighted Sum Energy Efficiency with imperfect CSI Considerations

Transmission errors are contributed by two factors, namely, channel noise and channel outage. Channel noise is usually caused by nonideal channel coding and/or finite block length of the channel, which can be normally addressed using strong coding (e.g., turbo code) and/or longer OFDM frame length. On the other hand, channel outage is a systematic effect caused whenever scheduled data rate exceeds maximum capacity, i.e., $r_{ij} > c_{ij}$, due to inaccurate realisations of actual channel gains $h = \{h_{ij}\}$. We calculate the channel outage probability by initially letting index $I\left[r_{ij} \leq c_{ij}\right] = \begin{cases} 1, & \text{if } r_{ij} \leq c_{ij} \\ 0, & \text{if } r_{ij} > c_{ij} \end{cases}$ to denote the throughput–capacity violation in binary form. The channel outage probability can be then accounted as the complementary probability of error-free transmissions conditioned on the estimated channel gain realisations $\hat{h} = \{\hat{h}_{ij}\}$, i.e., $P_{out,ij} = 1 - \Pr\left[r_{ij} \leq c_{ij} | \hat{h}\right]$. With the definition of $P_{out,ij}$ and recalling (3), the average WSEE (averaged over ergodic realisations of $\{h_{ij}\}$ and $\{\hat{h}_{ij}\}$) can be calculated as

$$
\begin{aligned}
\overline{EE} &= E_h\left[\text{WSEE}\right] \\
&= E_{\hat{h}}\left[E_{h|\hat{h}} \sum_{j=1}^{K} \sum_{i=1}^{N_F} \left(w_{ij} \cdot \frac{s_{ij} \cdot r_{ij} \cdot I[r_{ij} \leq c_{ij}]}{\zeta \cdot s_{ij} \cdot p_{ij} + P_{ij}^C}\right)\right] \\
&= E_{\hat{h}}\left[\sum_{j=1}^{K} \sum_{i=1}^{N_F} \left(w_{ij} \cdot \frac{s_{ij} \cdot r_{ij} \cdot E_{h|\hat{h}}\left[I[r_{ij} \leq c_{ij}]\right]}{\zeta \cdot s_{ij} \cdot p_{ij} + P_{ij}^C}\right)\right] \\
&= E_{\hat{h}}\left[\sum_{j=1}^{K} \sum_{i=1}^{N_F} \left(w_{ij} \cdot \frac{(1 - P_{out,ij}) \cdot s_{ij} \cdot r_{ij}}{\zeta \cdot s_{ij} \cdot p_{ij} + P_{ij}^C}\right)\right].
\end{aligned}
\tag{4}
$$

$\overline{EE}$ in (4) is random due to the unknown coefficients of actual channel gains $h = \{h_{ij}\}$ in $r_{ij} \leq BW \cdot \log_2\left(1 + \left(\eta \cdot p_{ij} \cdot |h_{ij}|^2\right)\right)$. To make $\overline{EE}$ tractable we correlate r_{ij} with $P_{out,ij}$ following a twofold rationality. Firstly, given any estimated channel gain $\hat{h}_{ij}$, the actual channel gain h_{ij} is Gaussian with mean $E_{h|\hat{h}}\left[h_{ij} | \hat{h}\right] = \hat{h}_{ij}$ and variance $E_{h|\hat{h}}\left[\left(h_{ij} - \hat{h}_{ij}\right) \cdot \left(h_{ij} - \hat{h}_{ij}\right) | \hat{h}\right] = \sigma_{\Delta h}^2$, i.e., $h_{ij} \sim \mathcal{CN}\left(\hat{h}_{ij}, \sigma_{\Delta h}^2\right)$. Therefore, term $\frac{|h_{ij}|^2}{\sigma_{\Delta h}^2}$ is non-central chi-squared random with two degrees of freedom and non-centrality parameter $\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}$. Secondly, we let $F_{x^2(\cdot)}(x)$ to denote the c.d.f. of non-central chi-squared random variables $x \in \mathbb{R}$ ($\mathbb{R}$ is the set of real numbers) with two degrees of freedom and $F_{x^2(\cdot)}^{-1}(x)$ its inverse c.d.f. Hence, we can calculate $P_{out,ij}$ as

$$
\begin{aligned}
P_{out,ij} &= 1 - \Pr\left[r_{ij} \leq c_{ij} | \hat{h}\right] \\
&= \Pr\left[r_{ij} > BW \cdot \log_2\left(1 + \eta \cdot p_{ij} \cdot |h_{ij}|^2\right) | \hat{h}\right] \\
&= \Pr\left[|\hat{h}_{ij}|^2 \leq \frac{2^{\frac{r_{ij}}{BW}} - 1}{\eta \cdot p_{ij}}\right].
\end{aligned}
\tag{5}
$$

From (5) we resolve

$$
r_{ij} = BW \cdot \log_2\left(1 + \eta \cdot \sigma_{\Delta h}^2 \cdot F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}(P_{out,ij}) \cdot p_{ij}\right),
$$

which correlates r_{ij} with $P_{out,ij}$. Then, we obtain the definition of $\overline{EE}$ in *Theorem 1* as follows.

Theorem 1. *Given the imperfect CSI realisation matrix $\hat{H}_{N_F \times K} = \left[\hat{h}_{ij}\right]$ the effective–weighted sum energy efficiency*

is defined as

$$\overline{EE} = E_{\hat{h}}\left[\sum_{j=1}^{K}\sum_{i=1}^{N_F} w_{ij} \cdot \frac{BW \cdot (1-P_{out,ij}) \cdot s_{ij} \cdot \log_2(1+\varphi_{ij} \cdot p_{ij})}{\zeta \cdot s_{ij} \cdot p_{ij} + P_{ij}^C}\right],$$

(6)

with the term $\varphi_{ij} = \eta \cdot \sigma_{\Delta h}^2 \cdot F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}(P_{out,ij})$ *to be added for synopsis.*

$\overline{EE}$ in *Theorem 1* is designed in a relatively simple manner accounting practical effects of data outage due to imperfect CSI using properties of inverse chi-squared c.d.f. Instead, existing efforts rely on exponential–type approximations of Marcum Q- and Bessel functions, which increase complexity and require recursive procedures to implement [23]–[26]. Also, our approach refers to the wide-ranging case of multi-user multi-carrier OFDMA transmissions to provide extensive perspectives of research and development in the green communication field. However, we comment that although similar interpretation of imperfect CSI has been applied to spectral efficiency problems in [4]–[5], [31]–[33], some key properties of inverse chi-squared c.d.f. still remain unexplored. For instance, even though literature provides evidences for the impact of $P_{out,ij}$ on $F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}(P_{out,ij})$ in (6), no study examines the impact of the non-centrality parameter $\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}$ neither on energy nor on spectral efficiency. Next sub-Section, provides such insights through inspecting the monotonicity of noncentral chi-squared inverse c.d.f. with respect to its noncentral parameter, which clarifies the impact of $\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}$ on $\overline{EE}$ and helps to present main properties of *Theorem 1*.

C. Properties of the Proposed Effective Weighted Sum Energy Efficiency

This sub–Section investigates the *user and channel selection principles* of $\overline{EE}$ aiming to show that the proposed objective in *Theorem 1* complies with fundamental properties of conventional ASEE. Particularly, it can be intuitively perceived by the analysis in [18] and [46] that with perfect CSI, ASEE in (2) is non-decreasing with respect to the actual channel gain realisation $h = \{h_{ij}\}$, i.e., larger $\{h_{ij}\}$ (better channel conditions) lead to higher ASEE. The challenge is that $\overline{EE}$ is a function of $\hat{h} = \{\hat{h}_{ij}\}$ (not $\{h_{ij}\}$ directly) and hence, we shall prove that with imperfect CSI and by given $\sigma_{\Delta h}^2$ and $P_{out,ij}$, $\overline{EE}$ is non-decreasing with respect to both actual and estimated channel gain realisations. In the following, we utilise the First-order logic theory [47] to introduce some lemma and theorem, which help to present the properties.

Lemma 1. *Let* $G(\chi,\theta)$, $\chi \in \mathbb{R}$ *denote a c.d.f. depending on a parameter* $\theta \in \Omega \subset \mathbb{R}$, *which is continuous and strictly increasing over* $\chi \in \{\chi \in \mathbb{R} : G(\chi,\theta) > 0\}$ *for each* $\theta \in \Omega$ *and non–increasing over* $\theta \in \Omega$ *for* $\chi \in \{\chi \in \mathbb{R} : G(\chi,\theta) > 0\}$. *Then the inverse* $G^{-1}(\psi,\theta)$, $\psi \in (0,1)$ *of* $G(\cdot,\theta)$ *is non–decreasing over* θ.

Theorem 2. *Let* $F(\chi,\theta)$, $\chi \in \mathbb{R}$ *denote the c.d.f. of a non–centralised chi–squared random variable with* κ *degrees of freedom and parameter* $\theta > 0$. *If* $F^{-1}(\psi,\theta)$, $\psi \in (0,1)$, *is*

the inverse of $F(\chi,\theta)$, $\chi \in (0,\infty)$, *then* $F^{-1}(\psi,\theta)$ *is non–decreasing with respect to* $\theta \in (0,\infty)$.

Proof. The proofs of *Lemma 1* and *Theorem 2* are presented in *Appendix A*. □

With *Lemma 1* and *Theorem 2* we investigate the monotonicity of the noncentral chi-squared inverse c.d.f. Within the analysis presented in *Appendix A* we show that the inverse function of a c.d.f. is non-decreasing not only over its random variable but also over its chi-squared non-centrality parameter. To our best knowledge, this aspect was not investigated previously and offers a unique insight into the structural determinants of energy and spectral efficiency objectives as considered in, e.g., [4]–[26], [30], [31]–[33]. Moreover, *Property 1* and *Property 2* follow readily from *Lemma 1* and *Theorem 2*.

Property 1. *[Impact of the estimated channel coefficient* $|\hat{h}_{ij}|^2$ *on* $\overline{EE}$*] The* $\overline{EE}$ *objective in Theorem 1 is non–decreasing with the estimated channel coefficient* $|\hat{h}_{ij}|^2$.

Property 2. *[Impact of the channel error variance* $\sigma_{\Delta h}^2$ *on* $\overline{EE}$*] The* $\overline{EE}$ *objective in Theorem 1 is non-decreasing with respect to large error variance* $\sigma_{\Delta h}^2$, *while for small to intermediate* $\sigma_{\Delta h}^2$ *has small increase ratio.*

Proof. The proofs of *Property 1* and *Property 2* are presented in *Appendix A*. □

Property 1 coincides with intuition. From it, $\overline{EE}$ will increase when the realisation of imperfect channel gain $\hat{h}_{ij}$ is more accurate and the actual channel gain h_{ij} is large. In physical terms, choosing the users with better channel quality helps to improve the overall energy efficiency like in ASEE. *Property 2* specifies that as the variance $\sigma_{\Delta h}^2$ of actual channel gain $h_{ij} \sim \mathcal{CN}(\hat{h}_{ij}, \sigma_{\Delta h}^2)$ increases, $\overline{EE}$ also increases. This is because it is well-known in physics that energy is proportional to the amplitude variations, which in our case is the variance $\sigma_{\Delta h}^2$. The two properties reveal that fundamental *user and channel selection* principles of the proposed objective in *Theorem 1* comply with the expected physics and mechanics of conventional ASEE in (2). Next Section adopts $\overline{EE}$ in (6) to formulate the energy efficiency maximisation problem with constraints on user QoS requirements, channel interference restrictions and transmitting power limitations.

IV. FORMULATION OF EFFECTIVE –WEIGHTED SUM ENERGY EFFICIENCY PROBLEM WITH IMPERFECT CSI

This section formulates the energy efficiency problem as a constrained maximisation problem and provides insights on its convexity considering subcarrier time–sharing relaxation.

A. Initial (Non–Convex) Optimisation Problem

The problem we consider is to jointly optimise subcarrier and power allocation so as to maximise WSEE of all users and subcarriers under imperfect CSI. Also, the problem aims to satisfy suppression of channel interference, a total transmitting power regulation and the individual QoS requirement of each user. In view of the $\overline{EE}$ definition in (6) and the system

modelling in Section II, the mathematical representation of the optimisation problem is as follows:

Find optimal policies $\mathcal{P}^* \left[\hat{H}_{N_F \times K} \right]$ *and* $\mathcal{S}^* \left[\hat{H}_{N_F \times K} \right]$

such that:
$$\max_{\mathcal{S}, \mathcal{P}} \overline{EE} \qquad \text{(bit/Joule)} \qquad (7)$$

subject to:
$$s_{ij} \in \{0, 1\}, \quad \forall i, j, \quad (8)$$

$$\sum_{j=1}^{K} s_{ij} = 1, \quad \forall i, \quad (9)$$

$$w_{ij} \geq 0, \quad \forall i, j, \quad (10)$$

$$p_{ij} \geq 0, \quad \forall i, j, \quad (11)$$

$$E \left[\sum_{j=1}^{K} \sum_{i=1}^{N_F} \left(s_{ij} \cdot p_{ij} \right) \right] \leq P_{TOTAL}, \quad (12)$$

$$E \left[\sum_{i=1}^{N_F} \left(\left(1 - P_{out,ij} \right) \cdot s_{ij} \times \right. \right.$$
$$\left. \left. BW \cdot \log_2 \left(1 + \varphi_{ij} \cdot p_{ij} \right) \right) \right] \geq q_j, \quad \forall j. \quad (13)$$

Constraint (8) ensures the interference avoidance among system subcarriers while, constraint (9) certifies the exclusive use of each subcarrier by one user only. Also, constraints (10) and (11) assure that weights and power allocations, respectively are non–negative, while (12) confirms the transmitting power feasibility through the upper power bound P_{TOTAL} for the BS. Finally, QoS constraint (13) guarantees the satisfaction of the minimum user throughput requirement q_j.

Problem (7)–(13) is mixed combinatorial because $\{p_{ij}\}$ are continuous and $\{s_{ij}\}$ are discrete variables. One solution is to perform an *exhaustive search* (ES) among all users and subcarriers such that for each subcarrier assignment, power is allocated to meet the individual user requirement and at the same time to maximise the energy efficiency [4], [6], [30], [36]. Although the optimal energy efficiency of the ES method, e.g., EE^{ES}, is the actual optimal solution of problem (7)–(13), it is well-known that ES method is inapplicable in reality due to its exceptionally high complexity of $\mathcal{O}\left(K^{N_F} \right)$.

B. Transformed (Convex) Optimisation Problem

To make the problem (7)–(13) more tractable we introduce the continuous variables $\tilde{s}_{ij} \in [0, 1]$ and $\tilde{p}_{ij} = p_{ij} \cdot \tilde{s}_{ij}$. The physical meaning of $\tilde{s}_{ij}$ is that a subcarrier can be shared in time among more than one user, while $\tilde{p}_{ij}$ indicates the corresponding power, which scales by $\tilde{s}_{ij}$. With $\tilde{s}_{ij}$ and $\tilde{p}_{ij}$ we can *transform* the subcarrier allocation constraints (8) and (9) from discrete to continuous, which *relaxes* the initial problem (7)–(13) into a potential convex problem. This relaxation technique is known as *subcarrier time–sharing* and has been extensively used to reduce complexity of spectral efficiency problems, where optimisation objectives are not in fractional forms, e.g., [4]–[5]. Remark that *subcarrier time–sharing* has not been applied in energy efficiency maximisation problems, where objectives are in fractional forms, to the best of our knowledge. Instead, relevant studies track solutions of problems similar to (7)–(13) through applying *fractional programming* as done in [7]–[26], where *subcarrier time–sharing* is not useful. For example,

with *fractional programming* the $\overline{EE}$ objective (7) could be modified to obtain the dual problem of problem (7)–(13) which however, would be likely to introduce uncontrollable duality gap and high complexity. By contrast, we use *subcarrier time–sharing* for first time to relax the subcarrier assignment constraint (8). We specify later that the difficulty in *subcarrier time–sharing* is that it results to transcendental equations, which require innovative analysis to derive feasible solutions.

Then, using the variables $\tilde{s}_{ij}$ and $\tilde{p}_{ij}$ we reformulate problem (7)–(13) as follows.

Find optimal policies $\tilde{\mathcal{P}}^* \left[\hat{H}_{N_F \times K} \right]$ *and* $\tilde{\mathcal{S}}^* \left[\hat{H}_{N_F \times K} \right]$

such that:
$$\max_{\mathcal{S}\left(\tilde{s}_{ij} \in [0,1], \sum_{j=1}^{K} \tilde{s}_{ij}=1 \right), \mathcal{P}\left(\tilde{p}_{ij} \geq 0 \right), w_{ij} \geq 0} \overline{EE} \qquad (14)$$

subject to:
$$E \left[\sum_{j=1}^{K} \sum_{i=1}^{N_F} \tilde{p}_{ij} \right] \leq P_{TOTAL}, \quad (15)$$

$$E \left[\sum_{i=1}^{N_F} \left(\left(1 - P_{out,ij} \right) \cdot \tilde{s}_{ij} \times \right. \right.$$
$$\left. \left. BW \cdot \log_2 \left(1 + \varphi_{ij} \cdot \frac{\tilde{p}_{ij}}{\tilde{s}_{ij}} \right) \right) \right] \geq q_j, \quad \forall j. \quad (16)$$

The convexity of the reformulated problem (14)–(16) is examined in *Proposition 1*.

Proposition 1. *The reformulated problem* (14)–(16) *is quasiconcave and determines the optimal solutions in the space* $(\tilde{s}_{ij}, \tilde{p}_{ij})$, *which is convex and non–empty.*

Proof. The proof of *Proposition 1* is presented in *Appendix B*. □

Proposition 1 verifies that the proposed energy efficiency problem (14)–(16) has same features to conventional dual problem in terms of convexity. Notice that for strictly quasiconcave and strictly concave functions, if a local maximum exists, it is also the global maximum [28], [47]. Therefore, the reformulated problem (14)–(16) has a unique global optimal solution, which always exists. Next section applies standard–Lagrangian optimisation to obtain the optimal solution of the reformulated problem (14)–(16) by means of final formulas.

V. JOINT OPTIMAL ALLOCATION SOLUTION BASED ON STANDARD–LAGRANGIAN OPTIMISATION

This Section presents a step-by-step analysis to obtain the joint optimal allocation solution of the proposed problem (14)–(16). We utilise standard–Lagrangian optimisation to apply an innovative mathematical method, which resolves the transcendental expressions issued due to the fractional nature of the $\overline{EE}$ objective (14). For the ease of presentation, we initially present the adaptive distribution of the optimal transmitting power and then the dynamic assignments of the optimal subcarrier indexing.

A. Optimal Transmitting Power Allocation Solution

The standard–Lagrangian function $\tilde{L}$ of the reformulated energy efficiency problem (14)–(16) is

$$\tilde{L} = \sum_{j=1}^{K} \sum_{i=1}^{N_F} \left(w_{ij} \cdot \frac{\left(BW \cdot (1 - P_{out,ij}) \cdot \tilde{s}_{ij} \cdot \log_2 \left(1 + \varphi_{ij} \cdot \frac{\tilde{p}_{ij}}{\tilde{s}_{ij}} \right) \right)}{\zeta \cdot \tilde{p}_{ij} + P^C_{ij}} \right) -$$

$$\mu \cdot \left(\sum_{j=1}^{K} \sum_{i=1}^{N_F} \tilde{p}_{ij} - N_F \cdot P_{TOT} \right) - \sum_{i=1}^{N_F} \nu_i \cdot \left(\sum_{j=1}^{K} \tilde{s}_{ij} - 1 \right) +$$

$$\sum_{j=1}^{K} \xi_j \cdot \left(\sum_{i=1}^{N_F} \left((1 - P_{out,ij}) \cdot \tilde{s}_{ij} \times \right. \right.$$

$$\left. \left. BW \cdot \log_2 \left(1 + \varphi_{ij} \cdot \frac{\tilde{p}_{ij}}{\tilde{s}_{ij}} \right) \right) - q_j \right),$$

where $\mu \geq 0$, $\xi_j \geq 0$, $\forall j$, $\nu_i \geq 0$, $\forall i$ are the Lagrangian multipliers associated with power constraint (15), QoS constraint (16) and relaxed subcarrier constraint (9), i.e., $\sum_{j=1}^{K} \tilde{s}_{ij} = 1$, respectively. Recall that relevant work either perform optimisation using dual–Lagrangian functions (e.g. the dual version of $\tilde{L}$) or develop numerical searching methods such as the generic algorithm in [30]. To avoid the uncontrollable duality gap of dual solutions [28], [47] and the practical issues on accuracy, complexity and convergence of numerical methods we shall build our analysis on the standard–Lagrangian function $\tilde{L}$.

According to the KKT conditions [28], the partial derivative of $\tilde{L}$ over $\tilde{p}_{ij}$ yields the optimal instantaneous transmitting power $\tilde{p}^*_{ij}$, i.e., $\left. \frac{\partial \tilde{L}}{\partial \tilde{p}_{ij}} \right|_{(\tilde{p}_{ij}, \tilde{s}_{ij}, \nu_i, \mu, \xi_j) = (\tilde{p}^*_{ij}, \tilde{s}^*_{ij}, \nu^*_i, \mu^*, \xi^*_j)} = 0$. We perform some calculations to conclude:

$$\frac{\Xi \left(\tilde{p}^*_{ij} \right)}{\ln 2 \cdot \left(\zeta \cdot \tilde{p}^*_{ij} + P^C_{ij} \right)^2 \cdot \left(1 + \frac{\varphi_{ij} \cdot \tilde{p}^*_{ij}}{s^*_{ij}} \right)} = 0, \qquad (17)$$

where

$$\Xi \left(\tilde{p}^*_{ij} \right) = \underbrace{\begin{array}{c} -\ln 2 \cdot BW \cdot (1 - P_{out,ij}) \cdot \zeta \cdot w_{ij} \cdot s^*_{ij} \times \\ \left(1 + \frac{\varphi_{ij} \cdot \tilde{p}^*_{ij}}{\tilde{s}^*_{ij}} \right) \cdot \log_2 \left(1 + \frac{\varphi_{ij} \cdot \tilde{p}^*_{ij}}{\tilde{s}^*_{ij}} \right) \end{array}}_{\Upsilon \left(\tilde{p}^*_{ij} \right)} +$$

$$BW \cdot (1 - P_{out,ij}) \cdot \zeta \cdot w_{ij} \cdot \varphi_{ij} \cdot \tilde{p}^*_{ij} +$$
$$BW \cdot (1 - P_{out,ij}) \cdot w_{ij} \cdot P^C_{ij} \cdot \varphi_{ij} -$$
$$\ln 2 \cdot \mu^* \cdot \left(\zeta \cdot \tilde{p}^*_{ij} + P^C_{ij} \right)^2 \cdot \left(1 + \frac{\varphi_{ij} \cdot \tilde{p}^*_{ij}}{\tilde{s}^*_{ij}} \right) +$$
$$BW \cdot \xi^*_j \cdot \varphi_{ij} \cdot (1 - P_{out,ij}) \cdot \left(\zeta \cdot \tilde{p}^*_{ij} + P^C_{ij} \right)^2 .$$

Since the denominator item in (17) is positive, we examine the numerator item $\Xi \left(\tilde{p}^*_{ij} \right)$. Equation $\Xi \left(\tilde{p}^*_{ij} \right) = 0$ is transcendental with respect to $\tilde{p}^*_{ij}$ given that $\Upsilon \left(\tilde{p}^*_{ij} \right)$ is a transcendental function, which cannot be directly solved. To solve (17) over $\tilde{p}^*_{ij}$ we determine the polynomial expression of $\Xi \left(\tilde{p}^*_{ij} \right)$ using its Maclaurin series expansion [47].

Theorem 3. *The transcendental function $\Xi \left(\tilde{p}^*_{ij} \right) = 0$ in (17) resolves the optimal instantaneous transmitting power $\tilde{p}^*_{ij}$ as*

$$\tilde{p}^*_{ij} = \left(\frac{s^*_{ij}}{\eta \cdot \sigma^2_{\Delta h} \cdot F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma^2_{\Delta h}} \right)} (P_{out,ij})} \cdot \chi^*_{ij} \right)^+, \qquad (18)$$

where χ^*_{ij} is the radical solution of the cubic Maclaurin polynomial of $\Xi \left(\tilde{p}^*_{ij} \right) = 0$ given in (29) and $(x)^+ = \max(0, x)$.

Proof. The proof of *Theorem 3* is presented in *Appendix C*. □

With *Theorem 3* and the analysis in *Appendix C* we show that the optimal instantaneous transmitting power $\tilde{p}^*_{ij}$ in (18) is a non–dual solution, which is originated by the transcendental function (17). Having defined $\tilde{p}^*_{ij}$, we focus next on the corresponding dynamic assignments of the optimal subcarrier indexing.

B. Optimal Subcarrier Allocation Solution

With the relaxed subcarrier index $\tilde{s}^*_{ij}$, the KKT conditions yield that the partial derivative of the Lagrangian function $\tilde{L}$ with respect to $\tilde{s}_{ij}$ is $\left. \frac{\partial \tilde{L}}{\partial \tilde{s}_{ij}} \right|_{(\tilde{p}_{ij}, \tilde{s}_{ij}, \nu_i, \mu, \xi_j) = (\tilde{p}^*_{ij}, \tilde{s}^*_{ij}, \nu^*_i, \mu^*, \xi^*_j)} = \begin{cases} 0, & \text{if } 0 < \tilde{s}^*_{ij} < 1 \\ > 0, & \text{if } \tilde{s}^*_{ij} = 1 \end{cases}$. With some calculations we obtain:

$$BW \cdot (1 - P_{out,ij}) \times$$
$$\underbrace{\left(\frac{\varphi_{ij} \cdot w_{ij}}{\zeta \cdot \tilde{s}^*_{ij} \cdot \chi^*_{ij} + \varphi_{ij} \cdot P^C_{ij}} + \xi^*_j \right) \cdot \left(\frac{\ln \left(1 + \chi^*_{ij} \right)^{1 + \chi^*_{ij}} - \chi^*_{ij}}{1 + \chi^*_{ij}} \right)}_{\Omega_{ij}} -$$

$$\nu^*_i = \begin{cases} 0, & \text{if } 0 < \tilde{s}^*_{ij} < 1 \\ > 0, & \text{if } \tilde{s}^*_{ij} = 1 \end{cases}. \qquad (19)$$

Given $\chi^*_{ij} > 0$, function Ω_{ij} in (19) is always positive since $\ln \left(1 + \chi^*_{ij} \right)^{1 + \chi^*_{ij}} - \chi^*_{ij} > 0$. The interpretation of function Ω_{ij} is to classify the weight of each subcarrier index among the $N_F \times K$ allocation combinations. Thus, we can use Ω_{ij} to evaluate the spectrum usage advantage of each of the K users on the same subcarrier. Furthermore, recalling that s^*_{ij} is the relaxed subcarrier index $\tilde{s}^*_{ij} \in [0, 1]$, its explicit definition is redundant to the initial problem (7)–(13) because $\overline{EE}^*$ is maximised when one subcarrier is assigned to one user only, i.e., $s^*_{ij} \in \{0, 1\}$. Note that the convex relaxation assumption introduced an intermediate coefficient to resolve the evaluation function Ω_{ij}, which must now consider the binary $s^*_{ij} \in \{0, 1\}$ in order to ensure channel interference cancellation and at the same time to allocate the subcarriers to the users that achieve the maximum $\overline{EE}^*$ increase. This can be done through searching among the N_F system subcarriers for the user with the maximum spectrum usage advantage, provided that the relaxed coefficients are identical. Therefore, (19) resolves to $s^*_{ij} = 0$ if $\nu^*_i > \overline{\Omega}_{ij}$, and $s^*_{ij} = 1$ if $\nu^*_i < \overline{\Omega}_{ij}$, with $\overline{\Omega}_{ij}$ the discrete version of Ω_{ij} to identify the weight of each binary subcarrier index $s^*_{ij} \in \{0, 1\}$. Having defined $\tilde{s}^*_{ij}$ and $\tilde{p}^*_{ij}$ we can now readily obtain the joint optimal allocation solution.

C. Joint Optimal Transmitting Power and Subcarrier Allocation Solution

Theorem 4. *The optimal transmitting power allocation policy $\tilde{\mathcal{P}}^* \left[\hat{H}_{N_F \times K} \right] = \left[\tilde{p}^*_{ij} \right]$ of the energy efficiency problem (14)–(16) has individual matrix element the optimal instantaneous transmitting power in (18). The corresponding optimal subcarrier allocation policy $\tilde{\mathcal{S}}^* \left[\hat{H}_{N_F \times K} \right] = \left[s^*_{ij} \right]$ has individual matrix element the optimal subcarrier allocation index*

$$s^*_{ij} = \begin{cases} 1, & j = j^* \\ 0, & \text{otherwise} \end{cases}, \qquad (20)$$

where j^* notates the optimal user selected by the linear search

$$\text{For each } i \text{ within } 1:N_F$$
$$j^* = \arg\max_j \overline{\Omega}_{ij} \quad , \tag{21}$$

and $\overline{\Omega}_{ij}$ is the subcarrier evaluation function in (19).

The results from *Theorem 4* are substituted in WSEE objective (6) to determine the optimal effective–weighted sum energy efficiency $\overline{EE}^*$ of problem (7)–(13). The joint solution in *Theorem 4* is independent from duality-gap and subgradient searching processes. This enables resolving to schemes with significantly lower complexity than dual approaches and achieving higher optimality. Also, with the analysis in *Appendix C* we show that our joint solution attains near-optimal points by radicals, which result to rather simple optimisation algorithms with guaranteed convergence. Then, $\overline{EE}^*$ can be obtained after having derived the Lagrangian multipliers ξ_j^* and μ^*, which are associated with the QoS constraint (13) and the power constraint (12), respectively. Next Section presents main properties of the introduced $\overline{EE}^*$ along with the searching process for ξ_j^* and μ^*.

VI. PROPERTIES OF THE PROPOSED JOINT OPTIMAL ALLOCATION SOLUTION

This Section provides detailed mathematical proofs to clarify that the joint optimal solution in *Theorem 4* and the introduced $\overline{EE}^*$ comply with main properties of the general energy efficiency framework [7]–[26]. We study the global optimality, the impact of imperfect CSI, the effect of circuit power consumption, the scheduling implementation process, the practical feasibility and the complexity of our proposal in the following sub–sections.

A. Properties of the Proposed Joint Optimal Solution on Global Optimality and Practical Feasibility

Property 3. *[Global optimality of the joint optimal solution] The optimal transmitting power and subcarrier allocation policies $\tilde{\mathcal{P}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{p}_{ij}^*\right]$ and $\tilde{\mathcal{S}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{s}_{ij}^*\right]$, respectively in Theorem 3 converge towards global optimum points with sufficiently small deviation.*

Property 4. *[Impact of the estimated channel gain $\left|\hat{h}_{ij}\right|^2$ on $\overline{EE}^*$] Given $\tilde{\mathcal{P}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{p}_{ij}^*\right]$ and $\tilde{\mathcal{S}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{s}_{ij}^*\right]$, the optimal effective–weighted sum energy efficiency $\overline{EE}^*$ in (14) is non–decreasing with the estimated channel coefficient $\left|\hat{h}_{ij}\right|^2$.*

Property 5. *[Impact of the circuit power P_{ij}^C on $\overline{EE}^*$] Given $\tilde{\mathcal{P}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{p}_{ij}^*\right]$ and $\tilde{\mathcal{S}}^*\left[\hat{H}_{N_F \times K}\right] = \left[\tilde{s}_{ij}^*\right]$, the optimal effective–weighted sum energy efficiency $\overline{EE}^*$ in (14) strictly decreases with the circuit power P_{ij}^C.*

Property 6. *[Impact of the circuit power P_{ij}^C on $\tilde{p}_{ij}^*$] The optimal instantaneous transmitting power $\tilde{p}_{ij}^*$ in (18) increases strictly with the circuit power P_{ij}^C.*

Proof. The proof of *Property 3* is similar to the proof presented in our previous work [33] and has been omitted due to space limitations. The proofs of *Property 4*, *Property 5* and *Property 6* are presented in *Appendix D*. $\square$

Property 3 specifies scheduling feasibility and global optimality of the proposed joint optimal solution. It shows that *Theorem 4* provides solutions close enough to the global optimum point, with practically negligible difference, i.e., within a controllable maximum truncation error of the approximation in (28) (see *Appendix C*). Later in Section VII we demonstrate with simulations that our joint optimal solution has sufficiently higher optimality than the relative dual approaches in [18]–[19], [33], [48].

Property 4 signifies that scheduler tends to select channels with higher realisation of imperfect channel $\hat{h}_{ij}$ than others. This also matches the conclusion of *Property 1*. Another outcome of *Property 4* is that as $\sigma_{\Delta h}^2$ is related to $\hat{h} = \{\hat{h}_{ij}\}$, the impact of the error variance $\sigma_{\Delta h}^2$ on $\overline{EE}$ can be seen through examining the channel gain (or loss) of the system due to multiuser diversity. The main idea is based on the fact that the number of users that participate in the system plays significant role regarding the gain magnitude, which increases as more users exist since the scheduler selects the best user among a large pool of candidates according to the qualities of each channel. Recalling *Lemma 1*, factor $F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}(P_{out,ij})$ grows in the same rate as parameter $\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}$. Thus, ϕ_{ij} in *Theorem 1* does not affect the growth order of multiuser diversity gain of $\overline{EE}^*$, which scales by $\Theta\left((1 - \sigma_{\Delta h}^2) \cdot \log(K)\right)$ for large K (see [31] for example)[4]. Similarly, when $\sigma_{\Delta h}^2 \to 1$ (no CSI) and the growth of users is limited, e.g., $(1 - \sigma_{\Delta h}^2) \cdot \log(K) \to \eta$, $\eta > 0$, the multiuser diversity gain is given by $\Theta\left(\left(\frac{\sigma_{\Delta h}^2}{2}\right) \cdot F^{-1}_{\left(\frac{\eta}{\frac{\sigma_{\Delta h}^2}{2}}\right)}(P_{out,ij})\right)$ as shown in [49]. In general, for intermediate channel errors, the multiuser diversity gain of $\overline{EE}^*$ decreases linearly as $\sigma_{\Delta h}^2$ increases exponentially when more users K is needed to compensate for the penalty of poor channel quality $\sigma_{\Delta h}^2$.

Furthermore, the intuition behind *Property 5* is that as circuit power increases, the optimal effective–weighted sum energy efficiency decreases. Also, the main point of *Property 6* is that as circuit power increases, higher power should be allocated to achieve higher throughput such that each information bit can be transmitted faster and less circuit energy is consumed. With *Property 3 – Property 6* we can verify that the proposed joint optimal solution in *Theorem 4* complies with practical properties of the general energy efficiency framework by means of global optimality and practical feasibility. Next, we present the implementation process of the proposed scheduling scheme.

B. Implementation Process of the Joint Optimal Solution

For many practical scenarios, the scheduler's ability to converge towards the optimal points is a very important feature. The convergence of the proposed scheduler depends on the

---

[4] $a_k = \Theta(b_k)$ if $\limsup_{K\to\infty} = \frac{|a_k|}{|b_k|} < \infty$ and $\limsup_{K\to\infty} = \frac{|b_k|}{|a_k|} < \infty$

evaluation process of the $K+1$ optimal Lagrangian multipliers μ^* and $\{\xi_j^*\}$, which are associated with the power and QoS constraints (15) and (16), respectively. To obtain μ^* and $\{\xi_j^*\}$ we use the KKT conditions of the energy efficiency problem (14)–(16) [28]:

Primal feasibility $: \sum_{j=1}^{K} \sum_{i=1}^{N_F} \tilde{p}_{ij}^* - P_{TOTAL} \leq 0,$

Dual feasibility $: \xi_j^* \geq 0, \ \forall j,$

Complementary slackness $: \xi_j^* \cdot \big(\sum_{i=1}^{N_F} \big((1 - P_{out,ij}) \cdot \tilde{s}_{ij} \times$
$$BW \cdot \log_2 \big(1 + \chi_{ij}^* \big) \big) - q_j \big) = 0, \ \forall j.$$

From the KKT conditions we build the below system of equations:

$$\begin{cases} \mathcal{F}_j \left(\boldsymbol{\xi}^*, \mu^* \right) = \xi_j^* \cdot \big(\sum_{i=1}^{N_F} \big((1 - P_{out,ij}) \cdot s_{ij}^* \times \\ \qquad\qquad \log_2 \big(1 + \chi_{ij}^* \big) \big) - q_j \big) = 0, \ \forall j, \quad (22) \\ \mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right) = P_{TOTAL} - \sum_{j=1}^{K} \sum_{i=1}^{N_F} \tilde{p}_{ij}^* = 0. \end{cases}$$

where vector $\boldsymbol{\xi}^*$ indicates the set of the K optimal ξ_j^*s, i.e., $\boldsymbol{\xi}^* = \{\xi_1^*, \xi_2^*, ..., \xi_K^*\}$, $j = 1, ...K$. The equation system (22) is solvable over μ^* and $\{\xi_j^*\}$ because it includes $K+1$ equations with $K+1$ unknown variables. For instance, *Fig 2* presents our algorithm in flowchart and discusses how the Lagrangian multiplier searching process can be carried to solve (22).

Description of the Lagrangian multipliers searching process: The process aims to obtain the joint optimal allocation solution in *Theorem 4* considering the effects of imperfect CSI $\hat{H}_{N_F \times K} = [\hat{h}_{ij}]$, the minimum users' throughput requirements $\{q_j\}$ and the total available power supplied at the BS P_{TOTAL}. In the beginning, it sets initial values for μ and $\{\xi_j\}$ and uses (19) and *Theorem* 4 to obtain the first estimation of the $N_F \times K$ (potentially optimal) subcarrier allocations $\tilde{\mathcal{S}} \big[\hat{H}_{N_F \times K} \big] = [\tilde{s}_{ij}]$. In part **A**, the K multipliers $\{\xi_j\}$ are updated using the Secant *root-finding method* (RFM) [28] such that the QoS constraint (16) of each user j is satisfied. If resources are insufficient, i.e., $\mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right) < 0$, the process exits. Else if $\mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right) \geq 0$ an inner loop in part **B** is initialised to redistribute the remaining power by updating μ using the bisection RFM [28]. With the new μ, the loop adjusts $\{\xi_j\}$ according to (22) and obtains the new $\tilde{\mathcal{S}} \big[\hat{H}_{N_F \times K} \big] = [\tilde{s}_{ij}]$. The loop stops when all the available power is totally exploited, i.e., $\mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right) = 0$, and the QoS of each user is satisfied, i.e., $\mathcal{F} \left(\boldsymbol{\xi}^*, \mu^* \right) < \delta$.

The proposed process in *Fig.2* ensures the geometric interpretation of the energy efficiency problem (14)–(16) because it considers satisfying the power and QoS constraints (15) and (16), respectively based on the necessary and sufficient KKT optimality conditions. Recalling from sub–Section II.A that actual channel gains $\{h_{ij}\}$ (and thus $\{|h_{ij}|^2\}$) are i.i.d. for each user, it is also ensured that the subcarrier indexes s_{ij}^*s (20) would be either 1 or 0. Therefore, the proposed searching process will always converge towards $\tilde{p}_{ij}^*$ and s_{ij}^* with QoS guarantee and total exploitation of the total available power P_{TOTAL} supplied at the BS. In addition, the convergence of the RFM in part **A** (find $\{\xi_j^*\}$ for fixed μ^*) is ensured as a similar mechanism is used in [50], where its convergence is proven. Also, as $\mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right)$ in (22) is monotonic in μ^*, we find μ^* for fixed $\{\xi_j^*\}$ via the bisection RFM in part **B**, which also
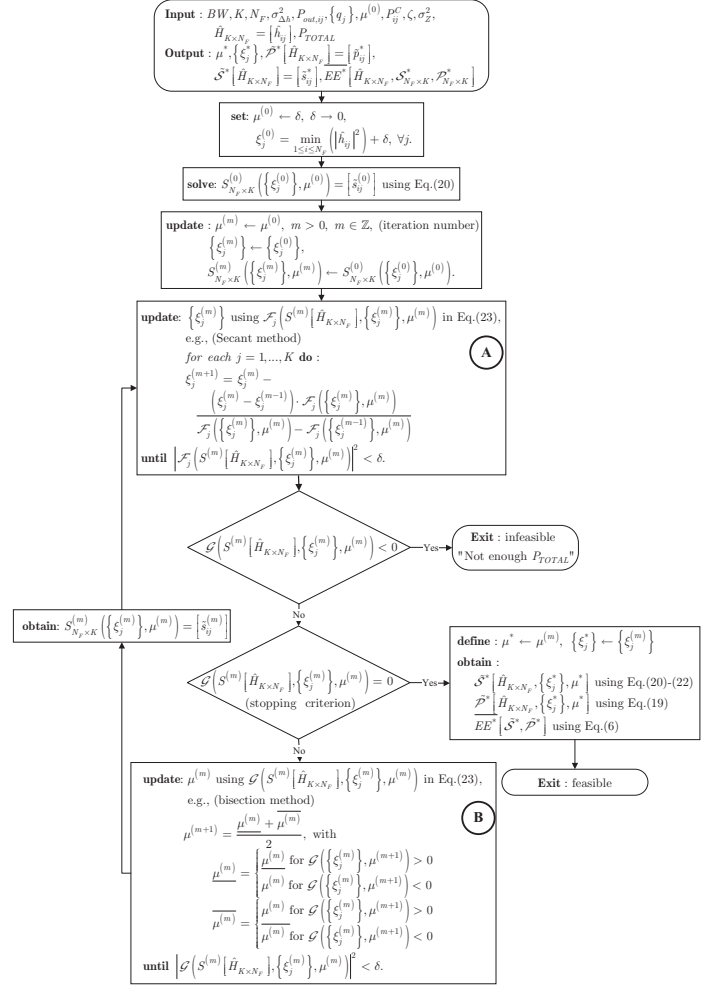

Fig. 2: Flowchart of the Lagrangian multiplier searching process in *Algorithm 1*.

converges once the initial range gives a bracketing interval. *Property 7* follows readily.

Property 7. *The joint optimal allocation solution in Theorem 4 satisfies the QoS and power constraints of the energy efficiency problem* (14)–(16) *and converges towards the global optimum points.*

C. Complexity of the Joint Optimal Allocation Solution

In continue, we examine the theoretical and practical complexity of the proposed joint optimal allocation solution. To obtain the subcarrier allocation policy, the implementation process defines the optimal user j^* through the linear search in (21). The search is on the number of users and subcarriers, which can be theoretically estimated by $\mathcal{O} \left(N_F \cdot K \right)$. Moreover, the search for the optimal multipliers μ^* and $\boldsymbol{\xi}^*$ lies in the number of iterations needed to update $\mathcal{F}_j \left(\boldsymbol{\xi}^*, \mu^* \right)$ and $\mathcal{G} \left(\boldsymbol{\xi}^*, \mu^* \right)$ in (22). Assuming that the RFMs have υ optimality, they require $\log_2 \left(\upsilon^{-1} \right)$ iterations to converge. Hence, the total computational complexity in practice becomes $\mathcal{O} \left(N_F \cdot K \cdot \log_2 \left(\upsilon^{-1} \right) \right)$. *Property 8* follows readily.

Property 8. *The total complexity of the joint optimal solution in Theorem 4 is of the order* $\mathcal{O}\left(N_F \cdot K \cdot \log_2\left(\upsilon^{-1}\right)\right)$.

The proposed scheme facilitates practical implementations as it has polynomial complexity with respect to the problem scale (N_F and K). The related study in [18] proposes a Dinkelbach's–type dual energy efficient algorithm with complexity $\mathcal{O}\left(N_F \cdot K \cdot (K+1)^2 \cdot \log_2\left(\upsilon^{-1}\right)\right)$, which is considerably higher than ours. The high complexity of [18] is due to the ellipsoid method that requires $\mathcal{O}\left((K+1)^2\right)$ time to obtain the $(K+1)$ dual variables. Some modifications of Dinkelbach's–type algorithm can be found in [19]–[22], which maximise energy efficiency with algorithmic complexity of $\mathcal{O}\left(I_{dual} \cdot 2 \cdot N_F \cdot K\right)$, where I_{dual} is the total number of inner iterations required for reaching convergence (e.g. [22], pp.7). Other relevant QoS-aware schemes based on dual–Lagrangian optimisation can be found in [4] and [48]. The algorithm in [4] maximises the throughput in $\mathcal{O}\left(K^3 \cdot \sqrt{N_F} \cdot \log_2\left(\upsilon^{-1}\right)\right)$, while power in [48] is minimised in $\mathcal{O}\left(N_F \cdot K^3 \cdot \log_2\left(\upsilon^{-1}\right)\right)$ time. As pointed out by the literature, the computational complexity of dual approaches is significantly higher than that of solutions based on convex optimisation.

VII. PERFORMANCE EVALUATIONS

For the fair comparisons, we adopt the similar simulation model and configurations in [18], [19] and [33], which maximise the energy efficiency subject to QoS constraints. The study in [18] proposes a *QoS-aware energy efficient resource scheduler* (QA-ERS), which differs from our scheme because it uses fractional programming to obtain dual optimal solutions under perfect CSI assumption. The algorithm in [19] uses *Dinkelbachs-based energy efficient joint allocation* (DEEJA), which has been employed finding an optimal solution similar to [18] at a lower complexity. Also, our previous efforts in [33] introduce another *Lagrangian-based error inconsiderate energy efficient* scheduler (SLEI–EE) assuming perfect CSI is availble. For performance comparison analysis, the simulation results of the dual-optimal throughput maximisation scheme [4] (maxRx) and the power minimisation scheme in [48] (minPx) have also been included. The system settings are given in *Table 1* considering *long term evolution* (LTE)–based links. Our simulations focus on average sum improvements of WSEE performance metric, which is the arithmetical mean of $\overline{EE}$ in (3) across all K users and N_F subcarriers by choosing $w_{ij} = (K \cdot N_F)^{-1}$, $\forall i,j$ in case of uniform weights. The impact of different weights is also shown in Section VII-D. For the ease of presentations we name our scheme as *joint subcarrier & power, QoS & imperfect channel–aware energy efficient* scheduler (JSPQ–IC–EE).

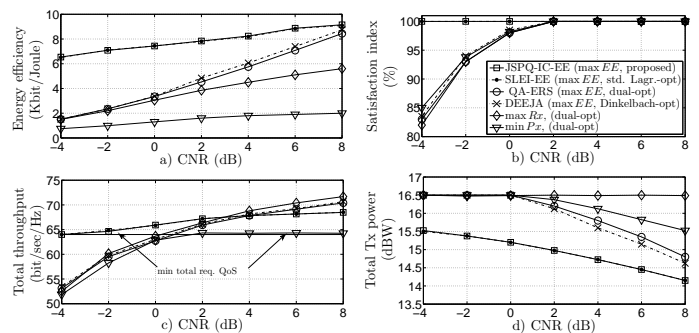### A. Evaluations under Perfect Channel Conditions

Fig.3(a) plots the energy efficiency vs. *channel–to–noise–rate* (CNR) considering perfect channel conditions, i.e., $P_{out,ij} = 0$, $\sigma^2_{\Delta h} = 10^{-8}$. We see JSPQ–IC–EE and SLEI–EE to outperform the other four schemes with significant gains in energy efficiency. Such performance gain mainly comes from the fact that the proposed solution method searches for the optimal points in a space larger than the space where the

TABLE 1. OFDMA SYSTEM SETTINGS FOR SIMULATION

Channel model	*6-tap Rayleigh frequency-selective*
Path loss (*same from BS to all users*)	*Hata urban propagation model* [52]
Channel estimation method	*nonsphericity–MMSE* [32]
Channel power delay profile parameter	0.1
Cell radius	0.5 Km
Noise spectral density (N_0)	$-174\ dBm/Hz$
Total channel bandwidth (BW)	$0.96\ MHz$
Total number of subcarriers (N_F)	64
Total number of users (K)	8
QoS of each user j (q_j)	$8\ bit/sec/Hz$
Channel outage probability ($P_{out,ij}$)	$0.01 - 0.3$
Channel error variance ($\sigma^2_{\Delta h}$)	$0.01 - 0.15$
Power inefficiency coefficient (ζ)	0.22
Total circuit consumption at BS (P^C)	$412.4\ W$
Circuit consumption per (i,j) link (P^C_{ij})	$\frac{P^C}{K \cdot N_F}\ W$
Total available Tx power (P_{TOTAL})	$45\ W$
Weight of each (i,j) link (w_{ij})	$\frac{1}{K \cdot N_F}$, or *varies according to Table 2*
CNR	$\{4, -2, 0, 2, 4, 6, 8\}\ dB,$

dual energy efficiency problem is determined. In JSPQ–IC–EE and SLEI–EE, the space is large because the schemes obtain $\overline{EE}^*$ imposing the total satisfaction for both power and QoS constraints. By contrast, QA-ERS and DEEJA determine a smaller space because dual solutions exploit a part of the available transmitting power, which means the QoS constraint in QA–ERS and DEEJA is weak (see [18], p.p 76–77 for example). Thus, more optimal points are excluded from the actual feasible space of the dual problem resulting to lower performance. It should be noted that Dinkelbach's–type approach in [19] enables DEEJA to achieve marginal gains over the conventional dual optimisation in QA-ERS. This reveals that the multiple parameterised subproblems in Dinkelbach's method can represent the actual feasible space more accurately than the conventional dual approach. On the other hand, maxRx and minPx seek within even smaller spaces, which correspond to the determination sets the QoS constraint (13) and the power constraint (12), respectively.

Furthermore, *Fig.3(b)* plots the *satisfaction index* (SI) of each scheme. SI is considered as a reliable metric for QoS performance and it is given by $SI = \frac{1}{K} \cdot \sum_{j=1}^{K}\left(\frac{\tilde{r}_j}{q_j}, 1\right)$, with $\tilde{r}_j$ the allocated rate to each user j. We see JSPQ–IC–EE and SLEI–EE guarantee users' QoS, while QA-ERS, DEEJA, maxRx and minPx cannot provide the required service when channel is noisy. This is because dual solutions cannot completely satisfy the QoS constraint (13) as allocation is done with power upper bound obtained at each scheduling slot, which may be either smaller or larger than the actual available power at the

Fig. 3: System performances vs. CNR under perfect channel conditions, i.e., $P_{out,ij} = 0, \sigma^2_{\Delta h} = 0$.
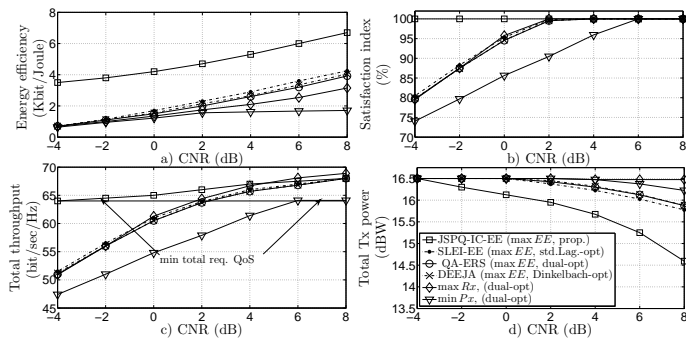
Fig. 4: System performances vs. CNR under imperfect channel, i.e., $P_{out,ij} = 0.1, \sigma^2_{\Delta h} = 0.01$.

Fig. 5: Energy–efficiency vs. channel error variance $\sigma^2_{\Delta h}$ and channel outage probability $P_{out,ij}$.

BS. The corresponding performances of total throughout and total transmitting (Tx) power are illustrated in *Fig.3(c)* and *Fig.3(d)*, respectively. From the graphs we can further observe the impact of the duality gap on system performances. Energy efficiency, QoS, throughput and power performances achieved by Lagrangian dual optimisation are lower than JSPQ–IC–EE and SLEI–EE because the duality gap does exist.

B. Evaluations under Imperfect Channel Conditions

Fig.4 plots system performances vs. CNR considering imperfect channel conditions[5], i.e., $P_{out,ij} = 0.1$, $\sigma^2_{\Delta h} = 0.01$. With the existence of imperfect CSI, all performances are degraded compared to the case of perfectly known CSI. Interestingly, JSPQ–IC–EE in *Fig.4(a)* provides substantial energy enhancements over error-inconsiderate schedulers. This is because our design is able to perceive the channel estimate $\left|\hat{h}_{ij}\right|^2$ and the probability of data outage of each subcarrier $P_{out,ij}$, which enables effective optimisation of energy efficiency. By contrast, error–inconsiderate designs treat the imperfect channel estimate $\left|\hat{h}_{ij}\right|^2$ as perfect (i.e. $|h_{ij}|^2$), which results to lower energy efficiency. We also observe SLEI–EE to slightly outperform QA–ERS and DEEJA, which reveals that upon applying our solution method on error–inconsiderate scheduling, more effective optimisations than Dinkelbachs method and conventional dual approaches can be obtained even under imperfect CSI conditions.

Furthermore, *Fig.4(b)* plots the corresponding SIs to examine the impact of imperfect CSI on QoS performances. The resilience of JSPQ–IC–EE to CSI errors enables total satisfaction of minimum QoS required by each user. As expected, the error–inconsiderate schedulers have lower QoS performances than JSPQ–IC–EE but not significantly lower than their corresponding QoS performances with perfect CSI in *Fig.3(b)*. For example, comparing *Fig.3(b)* with *Fig.4(b)* at $CNR = 2 \; dB$, we see JSPQ–IC–EE, SLEI–EE, QA–ERS, DEEJA satisfy all users but on the other hand their corresponding energy efficiencies in *Fig.3(a)* and *Fig.4(a)* degrade significantly by about $2 \; Kbit/Joule$. The phenomenon can be also observed in *Fig.4(c)* and *Fig.4(d)*, i.e., when resources are limited, the schemes prioritise consuming most of the available power (decrease energy efficiency) to obtain minimum
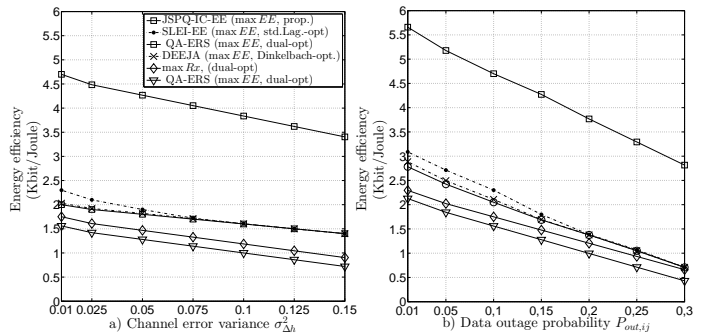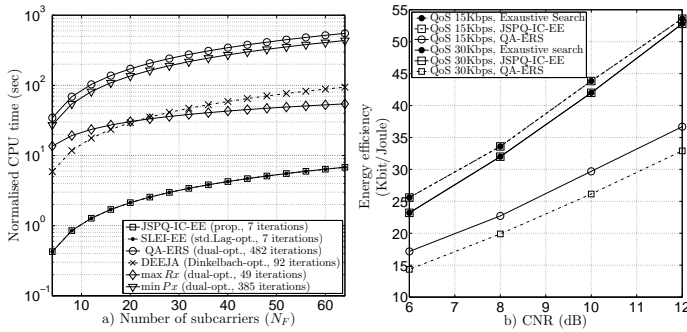
required throughput. In conclusion, the impact of imperfect CSI is greater on energy efficiency than on throughput and power performances.

To further examine the practical impact of CSI imperfectness on system performances, *Fig.5* plots the energy efficiency of each scheme with respect to the channel error variance $\sigma^2_{\Delta h}$ and outage probability $P_{out,ij}$. As expected, energy efficiencies decrease as either $\sigma^2_{\Delta h}$ or $P_{out,ij}$ increase. However, $\sigma^2_{\Delta h}$ in *Fig.5(a)* has less impact on energy efficiencies than $P_{out,ij}$ in *Fig.5(b)*. This is because energy in general is proportional to amplitude variations, which means increments of error variance can slightly increase the energy performance in an asymptotic manner. For this reason, all energy efficiencies in *Fig.5(a)* have lower reduction pace than energy efficiency in *Fig.5(b)*. Nevertheless, JSPQ–IC–EE brings substantially higher energy efficiency than the other schemes within all regions of $\sigma^2_{\Delta h}$ and $P_{out,ij}$.

C. Evaluations on Computational Time and Optimality

Fig.6(a) utilises the "PROFILE" function in MatLab to include crude average computational time measurements vs. the number of subcarriers. For the fair comparison, the optimality of all schemes is set to $\upsilon = 10^{-6}$. We see JSPQ–IC–EE and SLEI–EE need significantly lower computational time compared to all other examined schemes. This is because the optimal subcarrier assignments in (21) are decoupled between all N_F subcarriers (linear complexity with respect to N_F) and thus, the complexity of the optimal solution in *Theorem 4* is considerably low since it is depended only on the accuracy of the RFMs in part **A** and part **B** of *Fig. 2*. By contrast, the search for subgradients of each dual optimal subcarrier assignment in QA–ERS, DEEJA, maxRx and minPx scheduling imposes higher computational cost mainly due to the complexity of ellipsoid method. It should be noted that the decomposition of problem (7)–(13) into multiple parameterised problems in Dinkelbach's method can significantly reduce complexity of DEEJA compared to the conventional dual approach used in QA–ERS.

Fig.6(b) examines the accuracy of JSPQ–IC–EE and QA–ERS in relation to the exhaustive search under imperfect CSI conditions. The results of the exhaustive search are derived using the method in [28] and are considered as metrics of the theoretical optimality. For this experiment we set $\zeta = \frac{1}{0.38}$,

---

[5]The impact of channel imperfectness on error–inconsiderate schemes has been thoroughly studied in [32]–[31].

Fig. 6: *Computational complexity:* Normalised CPU time vs. number of subcarriers N_F. *Optimality:* Energy-efficiency vs. average CNR.
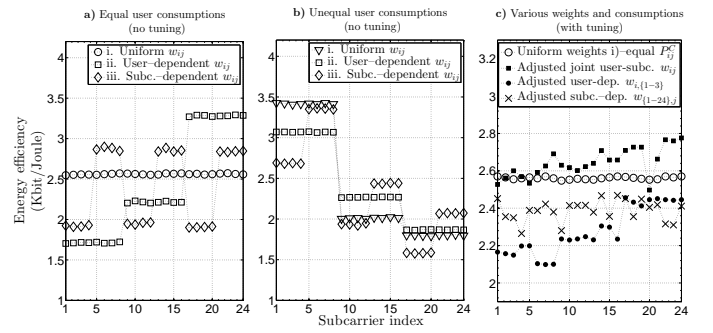

Fig. 7: Energy-efficiency vs. subcarrier index under a) equal user circuit power consumptions (no tuning), b) unequal user circuit power consumptions (no tuning) and c) unequal user circuit power consumptions (with tuning).

$P^C = 1W$, $P_{TOTAL} = 2W$ and we represent the imperfect CSI considering $P_{out,ij} = 0.1$ and $\sigma^2_{\Delta h} = 0.01$. We see JSPQ–IC–EE performs similarly to the exhaustive search, which means that firstly, our $\overline{EE^*}$ approximation in (28) is particularly accurate and secondly, the formative elements $P_{out,ij}$ and $\sigma^2_{\Delta h}$ of the imperfect CSI knowledge are correctly acknowledged by the proposed $\overline{EE}$ objective in *Theorem 1*. By contrast, QA–ERS deviates from the actual results by approximately 2 Kbit/Joule, which means the presence of imperfect CSI has considerable negative impacts on the accuracy of error-inconsiderate scheduling.

D. Evaluations under Various User and Subcarrier Weights

Fig.7 presents the impact of weights on the optimisation of JSPQ-IC-EE under perfect CSI ($P_{out,ij} = 0$, $\sigma^2_{\Delta h} = 0$) and $CNR = 2dB$. For this experiment, we consider fixed allocations of $N_F = 24$ subcarriers to $K = 3$ users, i.e, subcarriers 1-8, 9-16, and 17-24 are for users 1, 2, and 3, respectively. In each sub-figure, we examine three different distribution scenarios: i. uniform weights ($w_{ij} = \frac{1}{K \cdot N_F}$, $\forall i, j$); ii. user–dependent weights ($w_{i1} = \frac{0.05}{K \cdot N_F}$, $w_{i2} = \frac{0.6}{K \cdot N_F}$, and $w_{i3} = \frac{2.35}{K \cdot N_F}$, $\forall i$); and iii. subcarrier–dependent weights ($w_{ij} = \frac{0.2}{K \cdot N_F}$ for $i = \{1, 2, 3, 4\}$, $\{9, 10, 11, 12\}$, $\{17, 18, 19, 20\}$ and $w_{ij} = \frac{1.8}{K \cdot N_F}$ for $i = \{5, 6, 7, 8\}$, $\{13, 14, 15, 16\}$, $\{21, 22, 23, 24\}$, $\forall j$).

In *Fig.7(a)* we assume that users have equal circuit power consumptions, i.e., $P^C_{ij} = \frac{P^C}{K \cdot N_F}$. As expected, with uniform weights (distribution i.) the individual energy efficiencies are approximately similar, whereas non-uniform distributions ii. and iii. allow prioritising the energy efficiencies to specific users and subcarriers, respectively. More precisely, the user–dependent weight distribution ii. achieves highest energy efficiencies on subcarriers 17-24 (for user 3 with largest weight), whereas smallest energy efficiencies are observed on subcarriers 1-8 (for user 1 with lowest weight). Similarly, subcarrier–dependent weight distribution iii. leads to highest energy efficiencies on subcarriers 1–4, 9–12 and 17–20 (with largest weights), whereas smallest energy efficiencies are achieved on subcarriers 5–8, 13–16 and 21–24 (with lowest weights). In conclusion, *Fig.7(a)* shows that with equal circuit power consumptions, the energy efficiencies are dictated by the weight of each user and/or subcarrier.

In *Fig.7(b)* we assume that users have unequal circuit power, i.e., $P^C_{i1} = \frac{0.25 \cdot P^C}{K \cdot N_F}$, $P^C_{i2} = \frac{0.5 \cdot P^C}{K \cdot N_F}$ and $P^C_{i3} = \frac{0.75 \cdot P^C}{K \cdot N_F}$.

Different from the previous results in *Fig.7(a)*, in this case although the hierarchy of weights (e.g. $w_{i1} > w_{i2} > w_{i3}$ or $w_{1-4,j} < w_{5-8,j}$) the energy efficiencies are dictated by the circuit power consumption of each user (e.g. $P^C_{i1} < P^C_{i2} < P^C_{i3}$). For instance, with uniform weights (option i.), highest energy efficiencies are achieved on subcarriers 1–8 (for user 1 with the best/smallest power consumption P^C_{i1}), where lowest energy efficiencies are attained on subcarriers 17–24 (for user 3 with the worst/highest power consumption P^C_{i3}). Similarly, with user–dependent weights (option ii.), highest energy efficiencies are achieved on subcarriers 1–8 for user 1 and lowest energy efficiencies on subcarriers 19–24 for user 3. Notice that when considering option ii., the spread of the individual energy efficiencies across users is reduced as compared to option i., where option iii. allows to further reduce the energy spread acquiring the highest efficiencies to the selected (groups of) subcarriers. This means that by assigning weights over users and/or subcarriers, the energy spent by the system for each individual (i, j) allocation link can be controlled and hence, the energy concentration or the energy dispersal can be controlled over the available spectrum. In conclusion, *Fig.7(b)* shows that the impact of circuit power consumption on network performance is greater than the impact of weights, whereas weights allow to control the spread of energies (entropy energy) over the available spectrum.

Let us now gain insights on how the control of entropy energy can affect network performance. For this experiment, we perform dynamic allocations of subcarriers (not fixed as previously) by taking into consideration the various circuit power consumptions, channel coefficients and weighting factors as presented in *Table 2*. Our aim is to establish allocations such that entropy is reduced. We take as benchmark performance metric the energy efficiency performances achieved by the uniform weight distribution i. in *Fig.7(a)*, where equal circuit power consumptions were considered. We initially observe that adjusting either user weights (i.e. $w_{i,\{1-3\}}$) or subcarrier weights (i.e. $w_{\{1-24\},j}$), the highest energy efficiencies are achieved on subcarriers with better physical coefficients than others. Notably, in both cases entropy is significantly reduced compared to the corresponding graphs in *Fig.7(b)*. Also, adjusting subcarrier weights lead to slightly higher performance than adjusting user weights. This is because in the former case

TABLE 2. WEIGHTS AND CHANNEL GAINS FOR FIG.7(C)

User-dep. weights $w_{i,\{1-3\}}$	Subc.-dep. weights $w_{\{1-24\},j}$	Joint user-subc. weights $w_{i,j}$	Norm. channel gains *(dBm)*
$w_{i,1} = 0.15$	$w_{1,j} = 0.0189$	$w_{1,1} = 0.0200$	$h_{1,1} = -6.5892$
$w_{i,2} = 0.35$	$w_{2,1} = 0.0604$	$w_{7,1} = 0.0467$	$h_{7,1} = 5.0291$
$w_{i,3} = 0.50$	$w_{3,j} = 0.0859$	$w_{8,1} = 0.0700$	$h_{8,1} = 8.5511$
	$w_{4,j} = 0.0058$	$w_{11,1} = 0.0233$	$h_{11,1} = -18.4022$
	$w_{5,j} = 0.0251$	$w_{12,1} = 0.0300$	$h_{12,1} = -3.7521$
	$w_{6,j} = 0.0438$	$w_{18,1} = 0.0400$	$h_{18,1} = 1.8156$
	$w_{7,j} = 0.0642$	$w_{21,1} = 0.0500$	$h_{21,1} = 5.6393$
	$w_{8,j} = 0.0678$	$w_{23,1} = 0.0533$	$h_{23,1} = 6.1848$
	$w_{9,j} = 0.0287$	$w_{3,2} = 0.0333$	$h_{3,2} = -2.4118$
	$w_{10,j} = 0.0259$	$w_{4,2} = 0.0267$	$h_{3,2} = -3.4383$
	$w_{11,j} = 0.0321$	$w_{9,2} = 0.0500$	$h_{9,2} = -1.2922$
	$w_{12,j} = 0.0409$	$w_{10,2} = 0.0300$	$h_{10,2} = 1.1305$
	$w_{13,j} = 0.0496$	$w_{15,2} = 0.0433$	$h_{15,2} = 3.0591$
	$w_{14,j} = 0.0463$	$w_{16,2} = 0.0500$	$h_{16,2} = 2.3706$
	$w_{15,j} = 0.0467$	$w_{19,2} = 0.0467$	$h_{19,2} = 2.4567$
	$w_{16,j} = 0.0484$	$w_{24,2} = 0.0533$	$h_{24,2} = 2.8142$
	$w_{17,j} = 0.0221$	$w_{2,3} = 0.0300$	$h_{2,3} = -5.0250$
	$w_{18,j} = 0.0312$	$w_{5,3} = 0.0400$	$h_{5,3} = -1.5766$
	$w_{19,j} = 0.0280$	$w_{6,3} = 0.0500$	$h_{6,3} = -2.6587$
	$w_{20,j} = 0.0397$	$w_{13,3} = 0.0333$	$h_{13,3} = 0.8327$
	$w_{21,j} = 0.0559$	$w_{14,3} = 0.0367$	$h_{14,3} = 4.2549$
	$w_{22,j} = 0.0600$	$w_{17,3} = 0.0400$	$h_{17,3} = 4.9627$
	$w_{23,j} = 0.0341$	$w_{20,3} = 0.0600$	$h_{20,3} = -0.6878$
	$w_{24,j} = 0.0383$	$w_{22,3} = 0.0433$	$h_{22,3} = 0.4730$

we adjust 24 tuning parameters, whereas in the latter case only 3 parameters meaning that the less entropy the higher performance, and vice versa, which matches with the Second Law of Thermodynamics in Physics. Furthermore, by adjusting user and subcarrier weights jointly we utilise the maximum possible tuning parameters. In this situation, the network performance is higher not only than the cases of independent user/subcarriers adjustments but also than the case of uniform weight distribution i., where circuit power consumptions are equal. In conclusion, *Fig.7(c)* shows that joint adjustments of user and subcarrier weights can significantly reduce the entropy energy over the available spectrum to increase system performance.

VIII. CONCLUSION

This work considered the impact of imperfect CSI to propose an energy efficient resource scheduling solution for downlink transmissions in multiuser OFDMA networks. Unlike existing approaches, a new figure of merit for energy efficiency was proposed to address practical effects of priority considering heterogeneity between allocation links and data outage due to channel errors. It was shown that the proposed objective owns the similar user and channel selection properties with conventional definition of energy efficiency and that it can be degenerated to the energy efficiency model with perfect CSI. Additionally, the presented analysis examined the monotonicity of the noncentral chi-squared inverse c.d.f. to offer a unique insight into the structural determinants of the proposed objective. Furthermore, the study considered QoS, power and channel interference constraints to formulate an energy efficiency maximisation problem, which was shown to be quasiconcave with respect to power operands and convex with respect to subcarrier indexes over a convex non–empty set. Optimisation was then employed to resolve a rather simple optimal solution, which is capable of allocating the available powers and subcarriers jointly. We show that the joint optimal solution has guaranteed convergence, low complexity and high optimality due to its contemporary solution methodology,

which combines standard–Lagrangian analysis with Maclaurin polynomial approximations to attain the near-optimal points by radicals. The study conducted detailed examination on key properties of the joint optimal solution to show that it complies with fundamental features of conventional energy efficiency frameworks. Finally, fair evaluations were presented through simulations, which adopted similar system settings with related approaches. Illustrations confirmed the theoretical findings presenting significantly increased performances of the proposed scheme over relevant schedulers by means of energy efficiency, system throughput, transmitting power, QoS provision, implementation complexity and solution's optimality. The inclusion of multiple antennas will be considered in our future work.

APPENDIX A
PROOFS OF LEMMA 1, THEOREM 2, PROPERTY 1 AND PROPERTY 2

Proof of Lemma 1: Assume that $\theta_1 < \theta_2$. Since the c.d.f. G is monotone and continuous with respect to χ, for arbitrary $\psi \in (0,1)$ then there is some χ such that $\psi = G(\chi, \theta_1)$. Consequently, from the definition of the inverse function we obtain that $G^{-1}(G(\chi, \theta_1), \theta_1) = \chi = G^{-1}(G(\chi, \theta_2), \theta_2)$. Furthermore, since G is non–increasing with respect to θ it stands that $G(\chi, \theta_2) \le G(\chi, \theta_1)$, $\theta_1 < \theta_2$. As it is well–known that $G^{-1}(\psi, \theta)$ strictly increases with respect to ψ, we obtain that $G^{-1}(G(\chi, \theta_2), \theta_2) \le G^{-1}(G(\chi, \theta_1), \theta_2)$. Therefore, we conclude that $G^{-1}(G(\chi, \theta_1), \theta_1) \le G^{-1}(G(\chi, \theta_1), \theta_2)$ and thus $G^{-1}(\psi, \theta_1) \le G^{-1}(\psi, \theta_2)$ for $\psi \in (0,1)$. This completes the proof of *Lemma 1*. □

Proof of Theorem 2: According to *Lemma 1* it is sufficient to show that $F(\chi, \theta)$ is non–increasing over θ. Suppose that X is a random variable that follows non–centralised chi–squared distribution with κ degrees of freedom and parameter θ, symbolically $X \sim X_\kappa^2(\theta)$. Then by definition $X = X_1^2 + \sum_{m=2}^{\kappa} X_m^2$, where variables X_m, $m = 1, ...\kappa$, are independent with $X_1 \sim \mathcal{N}(\sqrt{\theta}, 1)$ and $X_m \sim \mathcal{N}(0,1)$, $m = 2, ...\kappa$. Therefore, $F(\chi, \theta) = \Pr(X \le \chi) = E[\Pr(X \le \chi | \Psi)]$, where $\Psi = \sum_{m=2}^{\kappa} X_m^2$. Since $\Pr(X \le \chi | \Psi) = \Pr(X_1^2 + \Psi \le \chi | \Psi) = \Pr(X_1^2 + \psi \le \chi)\big|_{\psi=\Psi}$ we obtain that

$$F(\chi, \theta) = E[g(\Psi, \theta)], \tag{23}$$

with $g(\psi, \theta) = \Pr(X_1^2 + \psi \le \chi)$ and the random variable $\Psi = \sum_{m=2}^{\kappa} X_m^2$ to have a centralised chi–squared distribution with $\kappa - 1$ degrees of freedom (note: independent of θ). Moreover, $g(\psi, \theta) = \begin{cases} \Pr(X_1^2 \le \chi - \psi), & \psi \le \chi \\ 0, & \psi > \chi \end{cases}$. For the special case $\psi \le \chi$ it stands that $g(\psi, \theta) = \Pr(|X_1| \le \sqrt{\chi - \psi}) = \Pr(-\sqrt{\chi - \psi} - \sqrt{\theta} \le X_1 - \sqrt{\theta} \le \sqrt{\chi - \psi} - \sqrt{\theta})$. But $X_1 - \sqrt{\theta} \sim \mathcal{N}(0,1)$ and consequently if Φ is the c.d.f. of $\mathcal{N}(0,1)$, then for $\psi \le \chi$ we obtain that $g(\psi, \theta) = \Phi(\sqrt{\chi - \psi} - \sqrt{\theta}) - \Phi(-\sqrt{\chi - \psi} - \sqrt{\theta})$. Thus, $g(\psi, \theta) = \Phi(\sqrt{\chi - \psi} + \sqrt{\theta}) + \Phi(\sqrt{\chi - \psi} - \sqrt{\theta}) - 1$, $\psi \le \chi$ and so, if f is the density function of $\mathcal{N}(0,1)$ then

$$\frac{\partial g}{\partial \theta} = \frac{1}{2\varepsilon}(f(\alpha + \varepsilon) - f(\alpha - \varepsilon)), \tag{24}$$

where $\alpha = \sqrt{\chi - \psi}$ and $\varepsilon = \sqrt{\theta}$. For the reason that f is sym-

metrical, non–decreasing within $(-\infty, 0)$ and non–increasing within $(0, \infty)$, it is easy to verify that $f(\alpha + \varepsilon) - f(\alpha - \varepsilon) < 0$, which accounting (24) yields that function g is non–increasing with respect to θ. Then by invoking (23) and *Lemma 1* we can conclude that $F^{-1}(\psi, \theta)$ is non–decreasing with respect to $\theta \in (0, \infty)$. This completes the proof of *Theorem 2*. $\square$

Proof of Property 1: To prove *Property 1* it is sufficient to show that $\varphi_{ij} = \eta \cdot \sigma_{\Delta h}^2 \cdot F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}(P_{out,ij})$ is non–decreasing with respect to $|\hat{h}_{ij}|^2$. We set $\theta = \frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}$ and use *Lemma 1* and *Theorem 2* to finally show that term φ_{ij} and the optimal effective–weighted sum energy efficiency objective $\overline{EE}$ (6) are non–decreasing with respect to $|\hat{h}_{ij}|^2$. This completes the proof of *Property 1*. $\square$

Proof of Property 2: To prove *Property 2* it is sufficient to verify that term φ_{ij} is non-decreasing with respect to large $\sigma_{\Delta h}^2$. For simplicity of calculations, we shall focus on the monotonicity of $\frac{F^{-1}(\psi, \theta)}{\theta}$ over $\sigma_{\Delta h}^2$, which is the same with the monotonicity of the term φ_{ij}. Calculating the derivative of $\frac{F^{-1}(\psi, \theta)}{\theta}$ we find that (*details are omitted due to space limitations*)

$$\frac{\partial}{\partial \theta}\left[\frac{F^{-1}(\psi, \theta)}{\theta}\right] = \frac{-\theta \cdot \frac{\partial F(\chi, \theta)}{\partial \theta} - \chi \cdot \frac{\partial F(\chi, \theta)}{\partial \chi}}{\theta^2 \cdot \frac{\partial F(\chi, \theta)}{\partial \chi}}, \ \theta > 0, \ \chi > 0. \quad (25)$$

Using the Marcum Q-function, the nominator of (25) is obtained as

$$\frac{1}{2} \cdot \left[\int_{\sqrt{\chi}}^{\infty} u \cdot e^{-\frac{u^2+\theta}{2}} \cdot \left[u \cdot \sqrt{\theta} \cdot I_1\left(u \cdot \sqrt{\theta}\right) - \theta \cdot I_0\left(u \cdot \sqrt{\theta}\right)\right]du - \chi \cdot e^{-\frac{\chi+\theta}{2}} \cdot I_0\left(\chi \cdot \theta\right)\right], \quad (26)$$

where I_0 and I_1 are the modified Bessel functions of the first kind. From (25) and (26), it can be straightforward verified that if $\theta \to 0$ then $\frac{\partial}{\partial \theta}\left[\frac{F^{-1}(\psi, \theta)}{\theta}\right] \to -\infty$. Therefore, at a space in the form of $(0, \omega)$, with ω dependent on ψ, the function $\frac{F^{-1}(\psi, \theta)}{\theta}$ is non-increasing for small $\theta = |\hat{h}_{ij}|^2/\sigma_{\Delta h}^2$. So, $\frac{F^{-1}(\psi, \theta)}{\theta} \to 0$ and thus, the term φ_{ij} are non-decreasing with respect to large values of $\sigma_{\Delta h}^2$. We also verify that under certain conditions of $P_{out,ij}$ and $|\hat{h}_{ij}|^2$, term φ_{ij} has a very slow increase pace, which means that for small to intermediate $\sigma_{\Delta h}^2$ the monotonicity of $\overline{EE}$ objective (6) is not affected. This completes the proof of *Property 2*. $\square$

APPENDIX B
PROOF OF PROPOSITION 1

The function[6] $\overline{EE}(\tilde{S}, \tilde{P})$ in (14) is a sum of functions in the form of $T(\tilde{S}, \tilde{P}) = \frac{(1-P_{out,ij}) \cdot \tilde{S} \cdot \tilde{R}(\tilde{S}, \tilde{P})}{\zeta \cdot \tilde{P} + P^C}$. To prove the quasi-concavity of $\overline{EE}(\tilde{S}, \tilde{P})$ it is sufficient to show that $T(\tilde{S}, \tilde{P})$ is quasiconcave. Let $\tilde{\mathcal{R}}_1^*, \tilde{\mathcal{R}}_2^*, \tilde{\mathcal{R}}_3^*$ to denote the optimal rate vectors corresponding to the overall throughput $\tilde{\mathcal{R}}_1, \tilde{\mathcal{R}}_2, \tilde{\mathcal{R}}_3$, respectively, with $\tilde{\mathcal{R}}_1 < \tilde{\mathcal{R}}_2 < \tilde{\mathcal{R}}_3$. Let us express the rate vector $\tilde{\mathcal{R}}_2^*$ as $\tilde{\mathcal{R}}_2^* = \lambda \cdot \tilde{\mathcal{R}}_1^* + (1 - \lambda) \cdot \tilde{\mathcal{R}}_3^*$, where $\lambda = \frac{\mathcal{R}_3 - \mathcal{R}_2}{\mathcal{R}_3 - \mathcal{R}_1}$, $0 < \lambda < 1$. From [48], [29] is known that $\mathcal{P}^*(\mathcal{R})$ is strictly convex

[6]To avoid confusion we drop the index $[\hat{H}_{N_F \times K}]$ in the Appendices.

in $\mathcal{R}$, which means $\tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_2) < \lambda \cdot \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_1^*) + (1 - \lambda) \cdot \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_3^*)$. Then $\tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_2^*) < \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_2)$ because $\tilde{\mathcal{R}}_2^*$ is the optimal rate vector among all the rate vectors with a summation of $\tilde{\mathcal{R}}_2$. Hence, $\tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_2^*) < \lambda \cdot \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_1^*) + (1 - \lambda) \cdot \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}_3^*)$ meaning that for any given throughput $\tilde{\mathcal{R}}$ the minimum transmitting power $\tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}) = \tilde{\mathcal{P}}^*(\tilde{\mathcal{R}}^*)$ is strictly convex in $\tilde{\mathcal{R}}$. If the superlevel set of $T(\tilde{\mathcal{R}})$ is $\mathcal{S}_\alpha = (\tilde{\mathcal{R}} \geq \mathcal{Q} \mid T(\tilde{\mathcal{R}}) \geq \alpha, \alpha \in)$, with α any real number and $\mathcal{Q}$ the minimum QoS requirement, then when $\alpha \geq 0$, $\mathcal{S}_\alpha = (\tilde{\mathcal{R}} \geq \mathcal{Q} \mid \alpha \cdot \zeta \cdot \mathcal{P}_{min}^{\tilde{*}}(\tilde{\mathcal{R}}) + P^C \leq 0)$, with $\mathcal{P}_{min}^{\tilde{*}}(\tilde{\mathcal{R}})$ the minimum total transmitting power required for any $\tilde{\mathcal{R}} \geq \frac{\mathcal{Q}}{BW}$. Otherwise when $\alpha < 0$ no points exist on the superlevel set. Therefore, from the convexity of $\tilde{\mathcal{P}}^*(\tilde{\mathcal{R}})$, $\mathcal{S}_\alpha$ is strictly convex in $\tilde{\mathcal{R}}$. This means the numerator item in $T(\tilde{S}, \tilde{P})$ is strictly convex in $\mathcal{P}$ and hence, $T(\tilde{P})$ is strictly quasiconcave. Moreover, we can easy verify that $\frac{\partial^2 T}{\partial \tilde{S}^2} < 0$, thus $T(\tilde{S})$ is strictly concave in $\tilde{S}$ as any positive linear combination of concave functions is concave. Consequently, $T(\tilde{S}, \tilde{P})$ is quasiconcave meaning that $\overline{EE}(\tilde{S}, \tilde{P})$ in (14) is also quasiconcave.

Let us now focus on the convexity of system constraints. Constraint (15) is straightforward affine, while the QoS constraint (16) is a sum of functions in the form of $G = (1 - P_{out,ij}) \cdot \tilde{S} \cdot BW \cdot \log_2\left(1 + \frac{\tilde{P}}{\tilde{S}}\right)$. Evaluating the second partial derivatives of G with s_{ij} and $\tilde{p}_{ij}$ we find that they are negative, i.e., $\frac{\partial^2 G}{\partial \tilde{S}^2} < 0$, $\frac{\partial^2 G}{\partial \tilde{P}^2} < 0$, thus the QoS constraint (16) is concave as any positive linear combination of strictly concave functions is strictly concave.

Let us now examine the set (space) where problem (14)–(16) is determined. Each of the constraints determines a convex set. Hence, the set defined by all constraints is the intersection of convex sets, which is also convex [28], [47]. However, the convex set may not exist or it can be an empty set. To prove the existence of $(\tilde{s}_{ij}, \tilde{p}_{ij})$ we denote with S^1 the feasible set over $\tilde{S}$ that satisfies the subcarrier constraints $\tilde{s}_{ij} \in [0, 1]$ and $\sum_{j=1}^K \tilde{s}_{ij} = 1$. Also, S^2 is the feasible set over $\tilde{P}$ that satisfies the constraints (15) and (16). Then in the $(N_F \times K + N_F \times K)$ space $(\tilde{s}_{ij}, \tilde{p}_{ij})$ the subcarrier constraints verify a cylinder with base Σ_1. Similarly the power constraints verify another cylinder with base Σ_2. The constraints in the space $(\tilde{s}_{ij}, \tilde{p}_{ij})$ determine the intersection of the two cylinders $\Sigma_1 \cap \Sigma_2$, which is a non–empty convex set. Hence, the strictly quasiconcave $EE(\mathcal{P})$ and the strictly concave $EE(\mathcal{S})$ are determined in the convex set $(s_{ij}, \tilde{p}_{ij})$. Consequently, $(\tilde{s}_{ij}, \tilde{p}_{ij})$ is convex and always exists, while given the strict quasiconcavity and the strict concavity of $EE(\mathcal{P})$ and $EE(\mathcal{S})$, respectively, the local maxima $EE(\mathcal{P}^*)$ and $EE(\mathcal{S}^*)$ are also global maxima. This completes the proof of *Proposition 1*. $\square$

APPENDIX C
PROOF OF THEOREM 3

Let us introduce the variable $\chi_{ij}^* = \frac{\varphi_{ij} \cdot \tilde{p}_{ij}^*}{\tilde{s}_{ij}^*}, \chi_{ij}^* > 0$, to represent transcendental function $\Xi(\tilde{p}_{ij}^*)$ as

$$\Xi(\tilde{p}_{ij}^*) = \mathcal{T}(\chi_{ij}^*) + \mathcal{E}(\chi_{ij}^*) = 0, \quad (27)$$

with

$$\mathcal{T}\left(\chi_{ij}^*\right) = \underbrace{-\ln 2 \cdot BW \cdot \left(1 - P_{out,ij}\right) \cdot \zeta \cdot w_{ij}}_{\Phi} \times$$

$$\left(1 + \chi_{ij}^*\right) \cdot \log_2\left(1 + \chi_{ij}^*\right),$$

$$\mathcal{E}\left(\chi_{ij}^*\right) = \varepsilon_1 \cdot \left(\chi_{ij}^*\right)^3 + \varepsilon_2 \cdot \left(\chi_{ij}^*\right)^2 + \varepsilon_3 \cdot \chi_{ij}^* + \varepsilon_4,$$

$$\varepsilon_1 = -\frac{\ln 2 \cdot \zeta^2 \cdot \mu^* \cdot \tilde{s}_{ij}^*}{\varphi_{ij}^2},$$

$$\varepsilon_2 = \frac{\ln 2 \cdot \zeta \cdot \tilde{s}_{ij}^*}{\varphi_{ij}^2} \cdot \left(\frac{BW \cdot (1 - P_{out,ij}) \cdot \xi_j^* \cdot \zeta \cdot \varphi_{ij}}{\ln 2} - \frac{2 \cdot P_{ij}^C \cdot \mu^* \cdot \varphi_{ij}}{\tilde{s}_{ij}^*} - \zeta \cdot \mu^*\right),$$

$$\varepsilon_3 = \left((1 - P_{out,ij}) \cdot \zeta \cdot P_{ij}^C\right) \cdot \left(2 \cdot BW \cdot \xi_j^* + \frac{BW \cdot w_{ij}}{P_{ij}^C} - \frac{\ln 2 \cdot \mu^* \cdot P_{ij}^C}{(1 - P_{out,ij}) \cdot \zeta \cdot \tilde{s}_{ij}^*} - \frac{2 \cdot \ln 2 \cdot \mu^*}{(1 - P_{out,ij}) \cdot \varphi_{ij}}\right),$$

$$\varepsilon_4 = \frac{BW \cdot (1 - P_{out,ij}) \cdot \left(P_{ij}^C\right)^2}{\tilde{s}_{ij}^*} \times \left(\xi_j^* \cdot \varphi_{ij} + \frac{w_{ij} \cdot \varphi_{ij}}{P_{ij}^C} - \frac{\ln 2 \cdot \mu^*}{BW \cdot (1 - P_{out,ij})}\right).$$

The complex function in (27) is the function $\mathcal{T}\left(\chi_{ij}^*\right)$, while $\mathcal{E}\left(\chi_{ij}^*\right)$ consists a cubic polynomial. Equations in the form of (27) can be effectively solved by determining the polynomial expression of $\mathcal{T}\left(\chi_{ij}^*\right)$ using its Maclaurin series expansion at point $\chi_{ij}^* = 0$ [47]. Point $\chi_{ij}^* = 0$ is the lowest possible point within the region $[0, +\infty)$, which is the interception of the feasible regions where χ_{ij}^* and $\tilde{p}_{ij}^*$ are determined, i.e., $\chi_{ij}^* \in (-1, +\infty)$ and $\tilde{p}_{ij}^* \in [0, +\infty)$. Consequently, the Maclaurin expansion will converge to the lowest possible power $\tilde{p}_{ij}^*$, which can asymptotically increase energy efficiency. The infinite Maclaurin series of polynomials of $\mathcal{T}\left(\chi_{ij}^*\right)$ is $\mathcal{T}\left(\chi_{ij}^*\right) = \sum_{n=0}^{\infty}\left(\mathcal{T}^{(n)}(0) \cdot (\chi_{ij}^*)^{(n)}\right)$, with $\mathcal{T}^{(n)}(0)$ to denote the n-th order Maclaurin polynomial. To track the solutions of the infinite series we express $\mathcal{T}\left(\chi_{ij}^*\right)$ in terms of finite series, i.e., given that the highest order of the polynomial part in (27) is third order, $\mathcal{T}\left(\chi_{ij}^*\right)$ can be denoted as $\mathcal{T}\left(\chi_{ij}^*\right) = \sum_{n=0}^{3}\left(\mathcal{T}^{(n)}(0) \cdot (\chi_{ij}^*)^{(n)}\right) + T_3$, with T_3 the truncation error. According to [33], [47] the truncation error is particularly small and the transcendental equation (27) can be well approximated as:

$$\Xi\left(\tilde{p}_{ij}^*\right) \approx \sum_{n=0}^{3}\left(\mathcal{T}^{(n)}(0) \cdot (\chi_{ij}^*)^{(n)}\right) + \mathcal{E}\left(\chi_{ij}^*\right)$$
$$= \left(\varepsilon_1 + \frac{\Phi}{6 \cdot \ln 2}\right) \cdot (\chi_{ij}^*)^3 + \left(\varepsilon_2 - \frac{\Phi}{2 \cdot \ln 2}\right) \cdot (\chi_{ij}^*)^2 + \left(\varepsilon_3 - \frac{\Phi}{\ln 2}\right) \cdot \chi_{ij}^* + \varepsilon_4 = 0.$$
$$(28)$$

The algebraic representation of the transcendental equation $\Xi\left(\tilde{p}_{ij}^*\right) = 0$ in (28) satisfies a cubic equation with respect to χ_{ij}^*. Using Cardano's theorem [47] we obtain the radical solutions of (28) with respect to χ_{ij}^*, i.e.,

$$\chi_{ij}^* = \sqrt[3]{\tau_1 + \sqrt{\tau_2^2 + \tau_3^3}} + \sqrt[3]{\tau_1 - \sqrt{\tau_2^2 + \tau_3^3}} - \tau_4, \quad (29)$$

with $\tau_1 = \frac{\alpha_2 \cdot \alpha_3}{6 \cdot \alpha_1^2} - \frac{\alpha_4}{2 \cdot \alpha_1} - \frac{\alpha_2^3}{27 \cdot \alpha_1^3}$, $\tau_2 = \frac{\alpha_2 \cdot \alpha_3}{6 \cdot \alpha_1^2} - \frac{\alpha_4}{2 \cdot \alpha_1} - \frac{\alpha_2^3}{27 \cdot \alpha_1^3}$, $\tau_3 = \frac{\alpha_3}{3 \cdot \alpha_1} - \frac{\alpha_2^2}{9 \cdot \alpha_1^2}$, $\tau_4 = \frac{\alpha_2}{2 \cdot \alpha_1}$, $\alpha_1 = -\frac{\left(\varepsilon_2 - \frac{1}{2 \cdot \ln 2}\right)}{3 \cdot \left(\varepsilon_1 + \frac{1}{6 \cdot \ln 2}\right)}$, $\alpha_2 = \alpha_1^3 + \frac{\left(\varepsilon_2 - \frac{1}{2 \cdot \ln 2}\right) \cdot \left(\varepsilon_3 - \frac{1}{\ln 2}\right) - 3 \cdot \left(\varepsilon_1 + \frac{1}{6 \cdot \ln 2}\right) \cdot \varepsilon_4}{\left(6 \cdot \left(\varepsilon_1 + \frac{1}{6 \cdot \ln 2}\right)^2\right)}$ and $\alpha_3 = \frac{\varepsilon_3 - \frac{1}{\ln 2}}{3 \cdot \left(\varepsilon_1 + \frac{1}{6 \cdot \ln 2}\right)}$. Finally, recalling $\chi_{ij}^* = \frac{\varphi_{ij} \cdot \tilde{p}_{ij}^*}{\tilde{s}_{ij}^*}$ and $\varphi_{ij} = \eta \cdot \sigma_{\Delta h}^2 \cdot F^{-1}_{\left(\frac{|\hat{h}_{ij}|^2}{\sigma_{\Delta h}^2}\right)}\left(P_{out,ij}\right)$ in (6), we obtain the optimal instantaneous transmitting power $\tilde{p}_{ij}^*$. This completes the proof of *Theorem 3*. $\square$
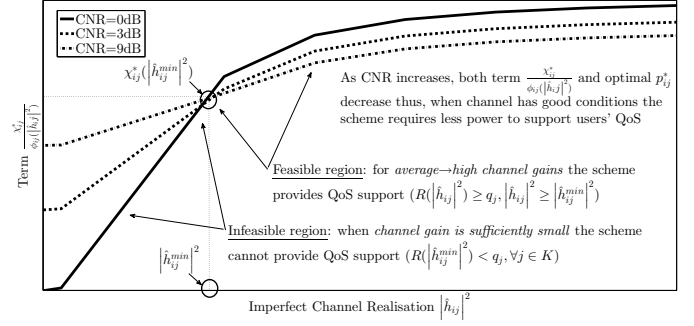
Fig. 8: *Specification on Property 4:* $\frac{\chi_{ij}^*}{\varphi_{ij}\left(\left|\hat{h}_{ij}\right|^2\right)}$–versus–$\left|\hat{h}_{ij}\right|^2$ curve for different levels of CNR.

APPENDIX D
PROOFS OF PROPERTY 4, PROPERTY 5 AND PROPERTY 6

Proof of Property 4: For brevity, let S, R, P and P_C to represent the subcarrier index, instantaneous data rate, radiated power and circuit power, respectively. The optimal energy efficiency is then defined in the fashion $EE^*(R) = \max_{\mathcal{S}}\left(\max_{\mathcal{P} \geq 0} EE(R, \mathcal{S})\right) = \max_{\mathcal{S}}\left(\sum \sum \frac{R}{\min_{\mathcal{P} \geq 0}(\zeta \cdot P(R, \mathcal{S}) + P_C)}\right)$. For any given $\mathcal{S}$, term $P(R, \mathcal{S})$ is minimised when power is distributed according to (18). Therefore, to examine the behaviour of EE in relationship with $\left|\hat{h}_{ij}\right|^2$ we should focus on term $\frac{\chi_{ij}^*}{\varphi_{ij}}$ of (18). The examination of χ_{ij}^* is performed through numerical representation in *Fig. 8*. In particular, term $\frac{\chi_{ij}^*}{\varphi_{ij}}$ increases as $\left|\hat{h}_{ij}\right|^2$ increases over certain CNR levels. Consequently if any $\left|\hat{h}_{ij}\right|^2$, $\min_{\mathcal{P} \geq 0} P(R, \mathcal{S})$ strictly decreases or remains the same. Then $\max_{\mathcal{P} \geq 0} EE(R, \mathcal{S})$ increases (but not necessarily strictly) with $\left|\hat{h}_{ij}\right|^2$. Thus, $EE(R, \mathcal{S})$ increases with $|h_{ij}|^2$, which means that the $EE^*(R)$–versus–R curve tends to be strictly higher or remains the same with the increase of $\left|\hat{h}_{ij}\right|^2$. Let us denote the new $EE^*(R)$ and R (with higher $\left|\hat{h}_{ij}\right|^2$) as $\overline{EE}^*(R)$ and $\overline{R}$, respectively. Then $EE^* \leq \overline{EE}^*\left(\overline{R}\right) \leq \overline{EE}^*$ meaning that the optimal EE^* increases with channel coefficient $\left|\hat{h}_{ij}\right|^2$. This completes the proof of *Property 4*. $\square$

Proof of Property 5: Furthermore, for any $\mathcal{P}$, $\mathcal{S}$, the optimal energy efficiency is a function in the form of $EE^*(\mathcal{R}) = \frac{\mathcal{R}(\mathcal{P}, \mathcal{S})}{\xi \cdot \mathcal{P}^*(\mathcal{R}(\mathcal{P}, \mathcal{S})) + P_C}$. Therefore, $EE^*(\mathcal{R})$ strictly decreases with P_C, which means the $EE^*(\mathcal{R})$–versus–$\mathcal{R}$ curve tends to be strictly lower with the increases of P_C. Let $\overline{EE}^*(\mathcal{R})$, $\overline{R}^*$ and $\overline{P_C}$ to denote the $EE^*(\mathcal{R})$, $\mathcal{R}^*$ and P_{ij}^C with larger circuit power. Then $\overline{EE}^* \leq EE^*\left(\overline{R}^*\right) \leq EE^*$ and thus, the optimal EE^* strictly decreases with P_{ij}^C. This completes the proof of *Property 5*. $\square$

Proof of Property 6: Define $\tilde{P}^*$ to be the optimal power allocation given a set of circuit power conditions $\left\{P_{C_j}\right\}$, with P_{C_j} the circuit power of user j. Also, let EE^* to represent the achieved energy efficiency. Suppose any $\left\{P_{C_j}\right\}$ decreases a

certain amount to $\left\{P_{C_j}\right\} - \Delta P_C$. Then from *Property 5*, EE^* will increase and (according to (18) in *Theorem 3*) $\tilde{P}^*$ will decrease strictly with energy efficiency EE^* meaning that $\tilde{P}^*$ increases strictly with the circuit power P_C. This completes the proof of *Property 6*. □

REFERENCES

[1] T. Edler and S. Lundberg, Energy efficiency enhancements in radio access networks, Ericsson Review, vol. 81, no. 1, pp. 42-51, 2004.

[2] R. Kumar and L. Mieritz, "Conceptualizing 'green' IT and data center-power and cooling issues," Gartner, Research Paper G00150322, Sep. 2007.

[3] Y. Chen, S. Zhang, S. Xu and G. Li, "Fundamental tradeoffs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, pp. 30–37, June 2011.

[4] M. Tao, Y.-C. Liang and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2190–2201, June 2008.

[5] D. Zhang, Y. Wang and J. Lu, "QoS aware relay selection and subcarrier allocation in cooperative OFDMA systems," *IEEE Commun. Lett.*, vol. 14, p. 294–296, Apr. 2010.

[6] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 28–35, Dec. 2011.

[7] X. Ge, X. Huang, Y. Wang, M. Chen, Q. Li, T. Han and C.-X. Wang, "Energy Efficiency Optimization for MIMO-OFDM Mobile Multimedia Communication Systems with QoS Constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, June 2014.

[8] X. Ge, J. Hu, C.-X. Wang, C. H. Youn, J. Zhang, and X. Yang, "Energy efficiency analysis of MISO-OFDM communication systems considering power and capacity constraints," *ACM Mob. Netw. Appl.*, vol. 17, no. 1, pp. 29-35, 2012.

[9] F. S. Haider, C.-X. Wang, H. Haas, E. Hepsaydir and X. Ge, "Energy-efficient subcarrier-and-bit allocation in multi-user OFDMA systems," *IEEE Conf. Veh. Technology (VTC-Spring)*, Yokohama, Japan, 2012.

[10] G. W. Miao, "Energy-efficient uplink multi-user MIMO," *IEEE Trans. Wireless Commun.*, vol. 12, no.5, pp. 2302–2313, May 2013.

[11] G. W. Miao, N. Himayat, G. Li and S. Talwar, "Low-complexity energy efficient scheduling for uplink OFDMA," *IEEE Trans. Commun.*, vol. 60, pp. 112–120, 2012.

[12] G. W. Miao, N. Himayat and G. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, pp. 545–554, 2010.

[13] K. Cheung, S. Yang and L. Hanzo, "Achieving maximum energy-efficiency in multi-relay OFDMA cellular networks: a fractional programming approach," *IEEE Trans. Commun.*, vol. 61, pp. 2746–2757, July 2013.

[14] I. Ku, C.-X. Wang, and J. S. Thompson, "Spectral-energy efficiency tradeoff in relay-aided cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4970–4982, Oct. 2013.

[15] C. Xiong, G. Li, S. Zhang, Y. Chen and S. Xu, "Energy - and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 3874–3886, Nov. 2011.

[16] O. Onireti, F. Heliot, and M. Imran, "On the energy efficiency-spectral efficiency trade-off of distributed MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3741–3753, Sep. 2013.

[17] O. Onireti, F. Heliot, and M. Imran, "On the energy efficiency-spectral efficiency trade-off in the uplink of CoMP system," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 556–561, Feb. 2012.

[18] X. Xiao, X. Tao and J. Lu, "QoS-aware energy-efficient radio resource scheduling in multi-user OFDMA systems," *IEEE Commun. Lett.*, vol. 17, pp. 75–78, Jan. 2013.

[19] Y. Wang, J. Zhang, P. Zhang, "Energy-efficient power and subcarrier allocation in multiuser OFDMA networks," *IEEE Int. Conf. Commun. ICC*, pp. 5492–5496, 10-14 June 2014.

[20] D. W. K. Ng, E.Lo, R. Schober, "Energy-Efficient Resource Allocation in Multi-Cell OFDMA Systems with Limited Backhaul Capacity," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3618 - 3631, Oct. 2012.

[21] R. S. Prabhu and B. Daneshrad, "An Energy-Efficient Water-Filling Algorithm for OFDM Systems," *IEEE Int. Conf. Commun. (ICC)* pp. 1–5, Cape Town, 2010.

[22] K. T. K. Cheung, S. Yang, L. Hanzo, "Achieving Maximum Energy-Efficiency in Multi-Relay OFDMA Cellular Networks: A Fractional Programming Approach," *IEEE Trans. Commun.*, vol.61, no.7, pp. 2746–2757, July 2013.

[23] X. Wang, F. Zheng, P. X. Zhu, "Energy-Efficient Resource Allocation in Coordinated Downlink Multi-cell OFDMA Systems," *IEEE Trans. Veh. Techn.*, doi: 10.1109/TVT.2015.2413950, 2015.

[24] E. Bedeer, O. Amin, O. A. Dobre, M. H. Ahmed, K. E. Baddour, "Energy-Efficient Power Loading for OFDM-Based Cognitive Radio Systems With Channel Uncertainties," *IEEE Trans. Vech. Technol.*, vol. 64, no. 6, pp. 2672–2677, June 2015.

[25] W. Xu, Y. Cui, H. Zhang, G. Y. Li, X. You, "Robust Beamforming With Partial Channel State Information for Energy Efficient Networks," *IEEE J. Sel. Areas Commun*, vol. 33, no. 12, pp. 2920–2935, Dec. 2015.

[26] L. Chen, Y. Yang, X. Chen, G. Wei, "Energy-Efficient Link Adaptation on Rayleigh Fading Channel for OSTBC MIMO System With Imperfect CSIT," *IEEE Trans. Veh. Techn.*, vol.62, no.4, pp. 1577–1585, May 2013.

[27] D. Palomar and M. Chiang, A tutorial on decomposition methods for network utility maximization, *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 14391451, Aug. 2006

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[29] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization for multicarrier systems," *IEEE Trans. Commun.*, vol. 54, pp. 1310-1322, July 2006.

[30] T. Ting, S. F. Chien, X.-S. Yang and S. Lee, "Analysis of quality-of-service aware orthogonal frequency division multiple access system considering energy efficiency," *IET J. Commun.*, vol. 8, pp. 1947–1954, 2014.

[31] Y. Yao and G. Giannakis, "Rate maximizing power allocation in OFDM based on partial channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1073–1083, May 2005.

[32] C. C. Zarakovitis, Q. Ni, D. Skordoulis and M. Hadjinicolaou, "Power-efficient cross-layer design for OFDMA systems with heterogeneous QoS, imperfect CSI, and outage considerations," *IEEE Trans. Veh. Technology*, vol. 61, pp. 781–798, Feb. 2012.

[33] C. C. Zarakovitis and Q. Ni, "Energy efficient designs for communication systems: resolutions on inverse resource allocation principles," *IEEE Commun. Lett.*, vol.17, no.12, pp. 2264–2267, Dec. 2013.

[34] J. G. Proakis, *Digital Communications*, New York: McGraw Hill, 2000.

[35] I. Isheden, G. P. Fettweis, "Energy-Efficient Multi-Carrier Link Adaptation with Sum Rate-Dependent Circuit Power," *IEEE Global Telecom. Conf. (GLOBECOM)*, pp. 1–6, Dec. 2010.

[36] C. Xiong, G. Li, Y. Liu, Y. Chen and S. Xu, "Energy-efficient design for downlink OFDMA with delay-sensitive traffic," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 3085–3095, June 2013.

[37] L. Venturino, C. Risi, S. Buzzi and A. Zappone, "Energy-efficient coordinated user scheduling and power control in downlink multi-cell OFDMA networks," in IEEE International Symposium Personal on Indoor and Mobile Radio Commun. (PIMRC), 2013.

[38] L. Venturino, C. Risi, A. Zappone and S. Buzzi, "Green joint user scheduling and power control in downlink multi-cell OFDMA networks," Future Network and Mobile Summit (Future Network Summit), 2013.

[39] L. Venturino, A. Zappone, C. Risi and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 1–14, Jan. 2015.

[40] I. Cosovic, S. Brandes and M. Schnell, "A technique for sidelobe suppression in OFDM systems," *IEEE Global Telecommun. Conf. GLOBECOM*, pp. 1–5, 2005.

[41] Y. Liu, L. Dong and R. J. Marks, "Joint reduction of out-of-band power and peak-to-average power ratio for non-contiguous OFDM systems," *IEEE Global Telecommun. Conf. GLOBECOM*, pp. 3493-3498, Atlanta, GA, 2013.

[42] Y. Ke, J. Craninckx, G. Gielen, A design approach for power optimized fully reconfigurable A/D converter for 4G radios, *IEEE Trans. on Circuits and Systems-II: Express Briefs*, vol. 55, no. 3, pp. 229-233, Mar. 2008.

[43] A. Suzuki, M. Inamori and Y. Sanada, "Subcarrier weighting scheme in OFDM receiver with sigma delta A/D converter on multipath fading channels," *IEEE Conf. Veh. Techn. Fall (VTC 2010-Fall)*, pp. 1–5, Ottawa, ON, 2010.

[44] A. Molisch and M. Win, MIMO systems with antenna selection," *IEEE Microwave Mag.*, vol. 5, pp. 46–56, Mar. 2004.

[45] Q. Ni, C. C. Zarakovitis, "Nash bargaining game theoretic scheduling for joint channel and power allocation in cognitive radio systems," *IEEE J. Sel. Areas Commun*, vol. 30, no. 1, pp. 70–81, Jan. 2012.

[46] S. Bayhan and F. Alagoz, "Scheduling in centralized cognitive radio networks for energy efficiency," *IEEE Trans. Veh. Technol.*, vol. 62, pp. 582–595, Feb. 2013.

[47] M. Greenberg, *Advanced Engineering Mathematics*, Prentice Hall, 1998.

[48] K. Seong, M. Mohseni and J. Cioffi, "Optimal Resource Allocation for OFDMA Downlink Systems," in *Proc. 2006 IEEE Int. Symp. Inf. Theory*, pp. 1394–1398.

[49] G. Song and Y. (G.) Li, Cross-layer Optimization for OFDM wireless networkPart I: Theoretical framework, *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614–624, Mar. 2005.

[50] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[51] C. C. Zarakovitis and Q. Ni, "A performance comparative study on the implementation methods for OFDMA cross-layer optimization," *Elsevier Future Generation Comput. Systems*, vol. 28, pp. 923-929, 2012.

[52] C. T. Ng and C. W. Sung, "Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1667-1675, 2008.

Charilaos C. Zarakovitis received the B.Sc. (2003) from Technical University of Crete, Greece, G.C.Eng. (2004) from Dublin Institute of Technology, Ireland, M.Phil. (2006) and Ph.D. (2012) degrees from Brunel University, London, U.K, all in electronic engineering. He is currently a Senior Research Associate at School of Computing and Communications, Lancaster University, Lancaster, U.K. He has successfully granted and carried out UK EPSRC Industrial CASE Awards and EU FP7 projects in partnership with Motorola Ltd., Nokia–Siemens–Networks Ltd., Intracom S.A., etc. His main research interests include next-generation green wireless networking, statistical signal processing, game theory and optimisation theory.

Qiang Ni received the B.Sc., M.Sc., and Ph.D. degrees from Huazhong University of Science and Technology (HUST), China, all in engineering. He is currently a Professor and the Head of the Communication Systems Group with InfoLab21, School of Computing and Communications, Lancaster University, Lancaster, U.K. He has published over 150 papers. His main research interests lie in the area of future generation communications and networking, including green communications and networking, cognitive radios, heterogeneous networks, 5G, energy harvesting, IoT and vehicular networks. He was an IEEE 802.11 Wireless Standard Working Group Voting Member and a Contributor to the IEEE Wireless Standards.

John Spiliotis received the B.Sc. (1974) from Kapodestrian University of Athens, Greece, DEA (1976) from Louis Pasteur University, Strasbourg, France and Ph.D. (1988) from National Technical University of Athens (NTUA), Greece, all in Mathematical Sciences. He is currently an Associate Professor at School of Applied Mathematics and Physical Sciences, NTUA, Athens, Greece. His main research interests are probability theory, stochastic analysis and stochastic control.