Zappone, A.; López-Pérez, D.; De Domenico, A.; Piovesan, N.; Bao, H. (2023). Rate, Power, and Energy Efficiency Trade-Offs in Massive MIMO Systems With Carrier Aggregation. IEEE Transactions on Green Communications and Networking. 7(3):1342-1355. https://doi.org/10.1109/TGCN.2023.3275302

Rate, Power, and Energy Efficiency trade-offs in Massive MIMO Systems with Carrier Aggregation

Alessio Zappone, David López-Pérez, Antonio De Domenico, Nicola Piovesan and Harvey Bao

***Abstract*—This work considers a multi-cell, multi-carrier massive MIMO network with carrier aggregation capabilities, and tackles both the rate versus power consumption and the rate versus energy efficiency (EE) trade-offs, by jointly optimizing the number of employed carriers, transmit antennas, base station density, and transmit power. Provably convergent algorithms for both trade-off problems are developed, together with closed-form results for the individual optimization of the considered resources, taking three main carrier aggregation techniques into account, namely inter-carrier, intra-carrier contiguous, and intra-carrier non-contiguous. Numerical results show how the use of carrier aggregation represents an effective way of increasing the network rate and EE, while keeping the power consumption at bay. By using carrier aggregation, it is possible to reduce the number of deployed antennas, without sacrificing the rate performance and increasing the system EE.**

I. Introduction

The rise of 5G wireless networks has brought unprecedented performance in terms of rate, ultra-reliable low latency and EE. Mobile network operators are able to offer innovative services to their customers, which brings us one step closer to the vision of a society with ubiquitous, high-capacity, reliable, and sustainable connectivity. Nevertheless, with 5G being a reality, the attention of the research community is focusing on how to address some of the limitations that mobile networks still have. In particular, one issue that deserves attention is the energy consumption of present and future wireless networks. Although the EE of 5G networks has improved compared to legacy 4G ones due to the significant increase in network throughput, the energy consumption level is a concern [1]. EE is explicitly mentioned as a key problem for future wireless networks by the Global System for Mobile communications Alliance (GSMA) [2] and the Next Generation Mobile Networks (NGMN) alliance [3], as it threats the environmental sustainability, and due to monetary reasons. In fact, the energy consumption constitutes a large portion of an operator's bill. In a recent report, it has been argued that, while 3GPP new radio (NR) deployments provide an EE level approximately four times higher than 3GPP long term evolution (LTE) ones, they also consume up to three times more energy, mainly due to the larger number of deployed antennas [4]. Moreover, the 4x increase of the EE observed is far from the 100x increase that the International Telecommunication Union (ITU) had set as objective for 5G [5]. At the same time, it has also forecasted that the number of internet protocol (IP) connections is going to annually rise by 55%, reaching 607 exabytes in 2025 and 5,016 exabytes in 2030, thus requiring an additional increase in the rate performance of future wireless networks. As a result, besides the use of the massive MIMO technology, it is of paramount importance to explore alternative approaches to provide the required rate levels, without causing a proportional increase of the energy consumption, to make future networks sustainable from an energy point of view.

A. Zappone is with the University of Cassino and Southern Lazio, Cassino, Italy, alessio.zappone@unicas.it. D. López-Pérez, A. De Domenico, N. Piovesan and H. Bao are with Huawei France R& D, Paris, France, {david.lopez.perez,antonio.de.domenico,nicola.piovesan,harvey.baohongqiang}@huawei.com

A. Carrier Aggregation

A technique with the potential of reaching the targeted data-rates with a sustainable power consumption is that of *carrier aggregation* (CA), as introduced by 3GPP NR. Specifically, CA is an enhanced multi-carrier operation, in which two or more component carriers are merged into *one* data channel, even if the component carriers are in separated frequency bands, e.g., sub-7GHz and millimeter wave spectrum. 3GPP NR supports the aggregation for a single mobile user (MS) up to 400 MHz in the sub-7GHz band and up to 700 MHz in the millimeter wave frequency region. The aggregation of component carriers can take place in three manners:

- Inter-band CA: component carriers from different frequency ranges are aggregated. This is the most flexible approach, but the hardest to implement, because component carriers from very different frequency bands may require a different transceiver chain due to the large frequency separation.
- Intra-band contiguous CA: adjacent component carriers from the same frequency range are aggregated. This is the least flexible approach, but the easiest to implement, because the aggregated channel can be considered as a single, enlarged channel from the radio frequency viewpoint, and thus a single transceiver chain may suffice.
- Intra-band non-contiguous CA: component carriers from the same frequency range, but not necessarily contiguous, can be aggregated. This approach represents a middle ground between the previous two, compromising between the flexibility of aggregation and the hardware requirements. The aggregated channel cannot be treated as a single, enlarged channel, thus requiring 1 transceiver chain per component carrier, but some degree of hardware sharing may still be possible among component carriers, e.g. multi-carrier filtering.

From the description above, it emerges that, just as it happens when increasing the number of antennas in a MIMO array, aggregating more component carriers will provide higher rates, but at the same time incur a larger power consumption, since more hardware is employed. This is true for all three aggregation techniques, but the hardware energy requirement will vary with the adopted carrier aggregation technique.

This work aims at answering the question whether it is more convenient to activate more component carriers or more antennas in a multi-cell network, which employs carrier aggregation and massive MIMO, from the point of view of increasing the rates, improving the EE, and lowering the power consumption. Specifically, the trade-off problems of rate versus power consumption and of rate versus EE will be considered with respect to the number of component carriers, active antennas, base station density, and transmit power.

B. State-of-the-art

Traditionally, the vast majority of works on the trade-off between rate, power, and EE focus on the optimization of the transmit power and/or the transmit beamforming, e.g., [6], [7], with few contributions focusing also on the design of the number of component carriers, number of antennas, and base station density.

In [8], a relay-assisted massive MIMO system is considered, and the trade-off between the spectral and EE is analyzed in the limit of a large number of relay antennas. It is shown that a large number of antennas can reduce multi-user interference, even though it increases the hardware power consumption. In [9], the spectral and the EE of a relay-assisted massive MIMO network are studied. It is shown that the EE function admits a unique maximizer with respect to both the number of antennas and the transmit power. In [10], the trade-off between the spectral and EE in heterogeneous massive MIMO networks is investigated. The problem is formulated, as a bi-objective program with the spectral and the EE as objectives, for the optimization of the user-cell associations, power and bandwidth allocations, and the number of active antennas. In [11], a general framework for the optimization of the EE of interference-limited wireless networks is developed by merging fractional programming and sequential optimization theory. In [12], the trade-off between the spectral and EE is analyzed in a massive MIMO system employing spatial modulation. The results indicate that, when a large number of transmit antennas is used, using a single radio frequency (RF) chain coupled with spatial modulation performs better than using multiple RF chains without spatial modulation. In [13], the trade-off between the spectral and EE is studied in a cell-free massive MIMO system considering the scenario in which the access points share information with a centralized controller by sending quantized versions of the signals. In [14], the spectral-energy Pareto-frontier is characterized in multi-cell MIMO networks resorting to the use of sequential fractional programming. In [15], the spectral and the energy efficiencies of a cell-free massive MIMO system are investigated. The weighted combination of these two metrics is optimized with respect to power control and to the association between MSs and access points. This leads to a mixed-integer problem, which is tackled by relaxing the discrete variables to continuous ones, and then employing the framework of sequential fractional programming.

Although all above works consider deterministic approaches in which the MS locations are estimated, a different line of research takes a stochastic approach, modeling the MS positions as random variables, and employing stochastic geometry to derive expressions of the performance metrics averaged over the MS positions. In this context, the vast majority of works embraces the so-called "standard modeling assumptions", among which there is the assumption of single-antenna transmission [16], [17], [18]. Few works have considered setups with multiple antennas. In [19], a massive MIMO system is considered, and a model that expresses the system power consumption as a function of the number of active antennas is introduced. The results show that the power consumption of a massive MIMO system grows at least linearly with the number of active antennas, and thus the EE is unimodal in the number of active antennas. These results are then extended in [20] to dense networks, where massive MIMO base stations are considered. In [21], the area spectral efficiency and the EE of a massive MIMO network are analysed, and the trade-off between them is also studied with respect to the number of active antennas and MSs. In [22], the trade-off between the energy and the spectral efficiencies is optimized with respect to the number of active antennas and the transmit power. In [23], a multi-cell massive MIMO system is considered, and the scaling laws of the EE with respect to the number of active antennas are derived. The results confirm that the EE of massive MIMO systems does not monotonically increase with the number of antennas. In [24], the scaling law of the EE is also studied with respect to a relay-assisted massive MIMO system. In [25], a dense massive MIMO network is considered, and the system EE is optimized with respect to the pilot reuse factor, the access point density, and the number of access point antennas and users. Finally, in [26], a full duplex, multi-carrier massive MIMO base station communicates with single antenna nodes, and the spectral and the energy efficiencies are optimized. These results are extended in [27] to the scenario in which the rate splitting approach coupled with successive interference cancellation is adopted.

All above works do not consider the CA technique. Indeed, the studies on the EE of CA are scarce, and have been focused mainly on legacy 4G networks. CA was shown to improve the performance of multi-carrier networks by allowing a base station to merge multiple physical carriers into a larger logical channel [28]. This clearly leads to a higher rates, but it can also increase the network energy consumption, with different energy consumption levels depending on its particular implementation [29]. In [30], a dynamic CA scheduling scheme is proposed, with the aim of improving the EE of uplink communications. Two scheduling methods are designed to reduce transmit power, while maximizing the utilization of wireless resources. In [31], the sum EE of a multi-carrier network is maximized with respect to the bandwidth and power allocation on each carrier component. A similar setup is considered in [32], and the problem of joint user-cell association and transmit power control for sum EE maximization is tackled. In [33], the bi-objective problem of power minimization and spectral efficiency maximization is considered in a system employing CA. In [34], the trade-off between EE and transmission delay in networks employing the CA technique is addressed. In [35], the EE of a multi-carrier system based on an heterogeneous network using CA is optimized with respect to the number of frequency bands, the active antennas, and the amount of feedback bits. In [36], an LTE-A-based network employing CA is considered, the optimization of an on-off operation of the transceiver is investigated. Finally, in [37], the optimization of the number of component carriers and antennas, as well as base station density and transmit power is tackled in a multi-carrier massive MIMO network, which a realistic power consumption model.

C. Contributions

At present, no research work considers the optimization of the rate versus power consumption or rate versus EE trade-offs in a system where CA and massive MIMO are jointly used. This work aims at filling this gap, making the following novel contributions.

- A stochastic geometry approach is taken, considering an average expression of the MS' rate for a multi-cell, multi-carrier, massive MIMO network. In this context, a power consumption model is derived, which accounts for the joint use of CA and massive MIMO, and relates the network power consumption to the base station density, MS density, and transmit power, as well as to the static power consumption of the whole system, which includes the power consumed for optimizing the radio resources.
- The rate versus power consumption and the rate versus EE trade-off problems are solved with respect to the number of component carriers, deployed antennas, base station density, and transmit power. Both problems are formulated as bi-objective optimization programs, for which the Pareto optimal frontier is computed.
- For each trade-off problem, the optimization of the different radio resources is separately investigated, and closed-form optimal allocation strategies are derived. These closed-form results represent the building blocks of the iterative algorithms that are proposed, ensuring a very limited overall computational complexity.
- Numerical results are provided to corroborate that carrier aggregation is a viable technique, which can be used together with massive MIMO to reduce the system power consumption without reducing the achieved rate, by trading off antennas with carriers. Indeed, our work shows that carrier aggregation may be a more energy efficient technique to achieve the rate requirements of MSs than massive MIMO.

The rest of the work is organized as follows. Section II describes the considered system model and problem formulation. Sections III and IV analyze the rate versus power consumption and the rate versus EE trade-off problems, respectively, providing the proposed solution algorithms. Finally, Section V presents a numerical performance analysis, while concluding remarks are provided in Section VI.

II. System Model and Problem Formulation

Let us consider the downlink of a multi-cell, multi-carrier, massive MIMO network operating in TDD mode. In this work, we take a stochastic geometry approach [38], in which the base stations are spatially distributed according to a homogeneous Poisson point process with intensity λ. The available communication bandwidth B is composed of N orthogonal component carriers managed by means of a CA technique. Inter-band CA, intra-band contiguous CA, and intra-band non-contiguous CA are all considered in this work. Moreover, we consider that each MS is allocated across all frequency resources of a cell, and that K single-antenna MSs are uniformly distributed within each Voronoi cell area. As a result, every base station has the same number of connected MSs. In addition, we consider that each base station is equipped with M transmit antennas and employs ZF precoding to remove intra-cell multi-user interference. User-cell association is performed based on the geographical distance and each MS is scheduled across all physical resource blocks (PRBs) of a component carrier. The path-loss between any base station and any given MS is modeled as ωd^{η}, with d the distance between the communicating nodes, η the path-loss exponent, and ω the path-loss at the reference distance. Flat fast fading is assumed, with fast-fading channels modeled according to the Rayleigh distribution, and considered independent among different base stations and MSs. Finally, the presence of hardware impairments is modeled by the presence of the impairment parameter $\epsilon \in [0,1]$, which causes both a decrease of useful signal power by a factor $1-\epsilon^2$ and the presence of an additional interference term with power ϵ^2 [19].

Adapting single-carrier model from [20] to the considered multi-carrier scenario, we consider that the base station transmit power p is uniformly split among MSs, component carriers, and MIMO layers, the average area rate (AR) (in $bit/s/m^2$) is,

$$AR = \lambda_{MS} N B \log_2\left(1 + \frac{(1-\epsilon^2)(M-K)}{\frac{2K}{\eta-2} + \epsilon^2(M-K) + \frac{\Gamma(1+\eta/2)}{\pi\lambda^2}\frac{\omega\,\sigma^2}{p}}\right), \tag{1}$$

wherein λ_{MS} is the mobile users' density, $\Gamma(\cdot)$ is the Gamma function, and σ^2 is the receive noise power. Thus, denoting by T_{opt} the time required to compute the radio resources to be employed, and by T the time duration before the radio resources are updated, the amount of information transmitted to the average MS is $(T-T_{opt})R$. On the other hand, the area energy consumption (AEC) (in $[J/m^2]$) during the complete time interval T is expressed as

$$AEC = \lambda\Bigg(NK(T-T_{opt})\mu p + \frac{N}{\beta_0}E_{FIX} + \frac{N}{\beta_1}E_{SYN} + \frac{N}{\beta_1}\mathcal{E}_0 M + N\mathcal{E}_1 MK + \mathcal{E}_2 K + E_{opt}\Bigg), \tag{2}$$

wherein

- $\mu \geqslant 1$ is the inverse of the transmit power amplifier efficiency, assumed to be independent of the transmit power.
- $E_{FIX} = P_{FIX}T$ is the load-independent energy consumed for control signaling, backhaul, cooling, and base band processing during the time T, with P_{FIX} being the

corresponding power consumption.

- $E_{SYN} = P_{SYN}T$ is the load-independent energy consumed by the local oscillators during the time T, with P_{SYN} being the corresponding power consumption.
- $\mathcal{E}_0$ is the overall energy consumed by each antenna transceiver during the time T.
- $\mathcal{E}_1 = f_cT/\eta_f$ is the overall energy consumed for signal processing purposes during the time T, wherein f_c is the clock frequency of the digital signal processor (and so f_cT is the total number of flops in a coherence block), while η_f is the computational efficiency in Watts/flop.
- $\mathcal{E}_2 = T\mathcal{A}\tau R_{BS}$ is the energy consumed for coding at the base station and for backhaul purposes during the time T, with τ the symbol interval, and $\mathcal{A}$ the coding power per MS.
- R_{BS} is the target data-rate of the base stations, while $E_{opt} = f_cT_{opt}/\eta_f$ is the energy consumed for radio resource optimization purposes.
- $\beta_0(N)$ and $\beta_1(N)$ model the type of carrier aggregation technique that is used. More in detail, in intra-band contiguous CA, the clock and the transceiver of each antenna can be shared among all the component carriers of a base station, which can be modeled by setting $\beta_0(N) = \beta_1(N) = N$ for all N. Instead, when inter-band CA is employed, each component carrier may need a different clock and a different transceiver per antenna, thus implying $\beta_0(N) = \beta_1(N) = 1$ for all N. By considering $\beta_0(N), \beta_1(N) \in \{1, \ldots, N\}$, every intermediate case can be modeled, e.g., the intra-band non-contiguous CA scenario. In any case, regardless of the specific carrier aggregation techniques that is used, physical considerations require that both functions β_0 and β_1 must increase at most linearly with N.

Thus, the bit-per-Joule network EE is defined as the amount of information transferred during the time T, over the corresponding energy consumption, as shown in (3), where it was accounted for the fact that, within the time T, information transfer takes place for $T - T_{opt}$ seconds. Moreover, the area power consumption (APC) (in $[W/m^2]$), defined as AEC$/T$, is shown in (4). with $\mathcal{P}_0 = \mathcal{E}_0/T$, $\mathcal{P}_1 = \mathcal{E}_1/T$, $\mathcal{P}_2 = \mathcal{E}_2/T$, and $P_{opt} = E_{opt}/T_{opt}$.

A. Problem formulation

The aim of this work is to analyze two trade-off problems, namely the rate versus power consumption trade-off and the rate versus EE trade-off, with respect to the optimization of the number of component carriers, active antennas, base station density, and transmit power. As for the number of MSs K per cell, let us observe that it is not a free parameter, since it is uniquely determined once λ_{MS} and λ are given. Indeed, if A is the service area, N_{BS} is the total number of base stations, and $K^{'} = KN_{BS}$ is the total number of MSs in the service area, then it holds that

$$K = \frac{KN_{BS}/A}{N_{BS}/A} = \frac{\lambda_{MS}}{\lambda} \ . \tag{5}$$

Plugging (5) into (1), (4), and (3), the trade-off problems to be solved can be formulated as bi-objective maximization programs, which, upon applying the scalarization technique, lead to

$$\begin{aligned}
&\max_{p,N,M,\lambda} \phi AR(p,N,M,\lambda) - (1-\phi)APC(p,N,M,\lambda) && \text{(6a)}\\
&\text{s.t. } AR(p,N,M,\lambda) \geqslant R_{min} && \text{(6b)}\\
&\quad P_{min} \leqslant p \leqslant P_{max} && \text{(6c)}\\
&\quad N_{min} \leqslant N \leqslant N_{max} && \text{(6d)}\\
&\quad M_{min} \leqslant M \leqslant M_{max} && \text{(6e)}\\
&\quad \lambda_{min} \leqslant \lambda \leqslant \lambda_{max} \ , && \text{(6f)}
\end{aligned}$$

for the rate versus power trade-off, and

$$\begin{aligned}
&\max_{p,N,M,\lambda} \phi AR(p,N,M,\lambda) + (1-\phi)\text{EE}(p,N,M,\lambda) && \text{(7a)}\\
&\text{s.t. } AR(p,N,M,\lambda) \geqslant R_{min} && \text{(7b)}\\
&\quad P_{min} \leqslant p \leqslant P_{max} && \text{(7c)}\\
&\quad N_{min} \leqslant N \leqslant N_{max} && \text{(7d)}\\
&\quad M_{min} \leqslant M \leqslant M_{max} && \text{(7e)}\\
&\quad \lambda_{min} \leqslant \lambda \leqslant \lambda_{max} \ , && \text{(7f)}
\end{aligned}$$

for the rate versus EE trade-off.

In both problems, R_{min} is the minimum MS's rate requirement, P_{max} and P_{min} are the maximum feasible and minimum required transmit power at the base station, N_{min} and N_{max} are the minimum and maximum number of component carriers, M_{min} and M_{max} are the minimum and maximum number of antennas, λ_{min} and λ_{max} are the minimum and maximum density of base stations.

In addition, in the problems, the boundary of the Pareto region can be described by varying $\phi \in [0, 1]$, with $\phi = 0$ and $\phi = 1$ corresponding to the two extreme points of the region associated to the individual optimization of the rate and power consumption for Problem (6) and of the rate and EE for Problem (7).

III. Area rate versus area power consumption trade-off

In order to tackle Problem (6) for any given value $\phi \in [0, 1]$, we resort to the alternating maximization algorithm, separately and iteratively optimizing N, M, λ, p. This choice is motivated by both complexity reasons and by the need of deriving some closed-form insight into the structure of the optimal solution. Indeed, Problem (6) is a mixed-integer problem, which is not convex with respect to the continuous optimization variables.

In the following three subsections, the optimization of N, M, λ, and p will be solved separately.

A. Optimization of N

The problem to be solved is stated as

$$\begin{aligned}
&\max_{N} \ F(N) && \text{(8a)}\\
&\text{s.t. } AR(N) \geqslant R_{min} && \text{(8b)}\\
&\quad N_{min} \leqslant N \leqslant N_{max} && \text{(8c)}
\end{aligned}$$

$$\text{EE} = \frac{(T-T_{opt})AR}{\text{AEC}} = \frac{NB\lambda_{MS}\log_2\left(1+\frac{(1-\epsilon^2)(M-K)}{\frac{2K}{\alpha-2}+\epsilon^2(M-K)+\frac{\Gamma(1+\alpha/2)}{\pi\lambda^2}\frac{\omega\sigma^2}{p}}\right)}{\lambda\Big(NK\mu p+\frac{T}{T-T_{opt}}\left(\frac{N}{\beta_0(N)}P_{FIX}+\frac{N}{\beta_1(N)}P_{SYN}+\frac{N}{\beta_1(N)}\mathcal{P}_0M+N\mathcal{P}_1MK+\mathcal{P}_2K\right)+\frac{T_{opt}}{T-T_{opt}}P_{opt}\Big)} \quad (3)$$

$$APC = \frac{AEC}{T} = \lambda\left(\frac{T-T_{opt}}{T}NK\mu p+\frac{N}{\beta_0(N)}P_{FIX}+\frac{N}{\beta_1(N)}P_{SYN}+\frac{N}{\beta_1(N)}\mathcal{P}_0M+N\mathcal{P}_1MK+\mathcal{P}_2K+\frac{T_{opt}}{T}P_{opt}\right) \quad (4)$$

with

$$\begin{aligned} F(N) &= \phi AR(N) - (1-\phi)APC(N) \\ &= \phi Na-(1-\phi)(\delta N+\psi g_0(N)+\gamma g_1(N)+\zeta)\ , \end{aligned} \quad (9)$$

and

$$a = B\lambda_{MS}\log_2\left(1+\frac{(1-\epsilon^2)(M-\lambda_{MS}/\lambda)}{\frac{2\lambda_{MS}/\lambda}{\eta-2}+\epsilon^2(M-\lambda_{MS}/\lambda)+\frac{\Gamma(1+\eta/2)}{\pi\lambda^2}\frac{\omega\sigma^2}{p}}\right) \quad (10)$$

$$\delta = \lambda_{MS}\mu p\frac{T-T_{opt}}{T}+\mathcal{P}_1M\lambda_{MS}\ ,\psi = P_{FIX}\lambda \quad (11)$$

$$\zeta = \mathcal{P}_2\lambda_{MS}+\lambda\frac{T_{opt}P_{opt}}{T}\ ,\gamma = \lambda(P_{SYN}+\mathcal{P}_0M), \quad (12)$$

and where we have also defined the functions

$$g_0(N) = \frac{N}{\beta_0(N)}\ ,\ g_1(N) = \frac{N}{\beta_1(N)}\ . \quad (13)$$

Clearly, the optimal number of component carriers N^* depends on the choice of the functions g_0 and g_1. However, in general the following result holds.

Proposition 1: For any β_0 and β_1 sub-linearly increasing with values in $[1, N]$, if $\phi a \geqslant (1-\phi)(\delta+\psi+\gamma)$, then the optimal number of component carrier is $N^* = N_{max}$. Instead, if $\phi a < (1-\phi)\delta$, then the optimal number of component carrier is

$$N = \max\left\{N_{min}, \left\lceil\frac{R_{min}}{a}\right\rceil\right\}\ . \quad (14)$$

Proof: Temporarily relaxing N to the continuous domain and studying the first-order derivative of F leads to

$$\frac{\partial F}{\partial N} \geqslant 0 \longleftrightarrow \phi a \geqslant (1-\phi)\left(\delta+\psi g_0^{'}(N)+\gamma g_1^{'}(N)\right)\ . \quad (15)$$

Next, let us observe that

$$g_0^{'}(N) = \frac{\beta_0(N)-N\beta_0^{'}(N)}{\beta_0^2(N)} \leqslant \frac{1}{\beta_0(N)} \leqslant 1 \quad (16)$$

$$g_1^{'}(N) = \frac{\beta_1(N)-N\beta_1^{'}(N)}{\beta_1^2(N)} \leqslant \frac{1}{\beta_1(N)} \leqslant 1\ , \quad (17)$$

because $\beta_0(N)$ and $\beta_1(N)$ are both increasing functions of N and are lower-bounded by 1 and upper-bounded by N. As a consequence, (15) is always implied by the condition

$$\phi a \geqslant (1-\phi)(\delta+\psi+\gamma)\ . \quad (18)$$

Thus, if (18) holds, the objective F is monotonically increasing with N, and hence the solution of Problem (8) is $N = N_{max}$.

Next, to show the second part of the thesis, let us recall that β_0 and β_1 are sub-linear functions, thus implying that $g_0^{'}(N) \geqslant 0$ and $g_1^{'}(N) \geqslant 0$. Therefore, if

$$\phi a < (1-\phi)\delta\ , \quad (19)$$

then the condition in (15) can never be fulfilled, regardless of the choice of the functions g_0 and g_1. In this case, the objective function F is monotonically decreasing, thus implying that the optimal solution of Problem (8) is the minimum feasible number of component carriers. Then, since the rate constraint is fulfilled by any $N \geqslant \left\lceil\frac{R_{min}}{a}\right\rceil$, we obtain the thesis. ■

Remark 1: Inspecting the condition in (18) we see that the right-hand-side corresponds to the inter-band CA case where $\beta_0(N) = \beta_1(N) = 1$ for all N, which therefore can be regarded as the worst-case as far as power consumption is concerned. Instead, the condition in (19) corresponds to the intra-band CA case where $\beta_0(N) = \beta_1(N) = N$ for all N, which therefore can be regarded as the best-case as far as power consumption is concerned.

Based on the remark above, we can provide an interpretation of the condition in (18). The left-hand-side depends on the priority of the rate ϕ and on the rate-per-carrier a, while the right-hand-side on the priority of the power consumption $1-\phi$ and on the worst-case power consumption per carrier. Thus, if the rate-per-carrier weighted by the priority assigned to the rate is higher than the worst-case power consumption per carrier weighted by the priority assigned to saving power, then the optimal strategy is to employ all available component carriers. If instead, the rate-per-carrier weighted by the priority assigned to the rate is lower than the best-case power consumption per carrier weighted by the priority assigned to saving power, then the optimal strategy is to employ the minimum feasible number of component carriers. Moreover, since these results hold for any choice of β_0 and β_1, they are true for any of the three carrier aggregation methods, i.e., inter-carrier, intra-carrier contiguous, and intra-carrier non contiguous.

If instead neither (18) nor (19) hold, i.e.,

$$(1-\phi)\delta \leqslant \phi a < (1-\phi)(\delta+\psi+\gamma)\ , \quad (20)$$

the optimal number of component carriers depends on the particular shape of the functions β_0 and β_1. For any choice of β_0 and β_1, it is always possible to solve (15) numerically to obtain the solution of the problem. Let us illustrate this in detail for the three notable cases of inter-band carrier aggregation, intra-band contiguous carrier aggregation, and intra-band non-contiguous carrier aggregation.

1) Inter-band carrier aggregation: Assume (20) holds, and set $\beta_0(N) = \beta_1(N) = 1$, which is a suitable model for inter-band carrier aggregation. Elaborating from (15) and plugging the expression of the derivatives of $g_0(N)$ and $g_1(N)$ leads to the condition

$$\phi a \geqslant (1-\phi)(\delta + \psi + \gamma) \ . \tag{21}$$

Thus, since (20) holds, the objective function F is monotonically decreasing in this case, and the optimal solution is the minimum feasible value of N, i.e. $N^* = \max\left\{N_{min}, \left\lceil \frac{R_{min}}{a} \right\rceil\right\}$.

2) Intra-band contiguous carrier aggregation: Assume (20), and set $\beta_0(N) = \beta_1(N) = N$, which is a suitable model for intra-band contiguous carrier aggregation. Elaborating from (15) and plugging the expression of the derivatives of $g_0(N)$ and $g_1(N)$ leads to the condition

$$\phi a \geqslant (1-\phi)\delta \ . \tag{22}$$

Thus, since (20) holds, the objective function F is monotonically increasing in this case, and the optimal solution is $N^* = N_{max}$.

3) Intra-band non-contiguous carrier aggregation: Assume (20), and set $\beta_0(N) = \beta_1(N) = \frac{N+1}{2}$, which is a suitable model for intra-band non-contiguous carrier aggregation. Elaborating from (15) and plugging the expression of the derivatives of g_0 and g_1 leads to the condition

$$(\phi a - (1-\phi)\delta)(N+1)^2 \geqslant 2(\psi+\gamma)(1-\phi) \ . \tag{23}$$

Due to (20), we have that $(\phi a - (1-\phi)\delta) > 0$, and thus we can find that

$$\bar{N} = \sqrt{\frac{2(\psi+\gamma)(1-\phi)}{\phi a - (1-\phi)\delta}} - 1 \ . \tag{24}$$

Moreover, (24) can be seen to be positive whenever $\phi a < (1-\phi)(\delta + 2(\psi+\gamma))$, which is true due to (20). Thus, when intra-carrier aggregation with non-contiguous bandwidth is used, the optimal number of component carriers is given by

$$N^* = \max\left\{\max\left\{N_{min}, \left\lceil \frac{R_{min}}{a} \right\rceil\right\}, \min\{\lceil \bar{N} \rceil, N_{max}\}\right\} \ , \tag{25}$$

where the constraints of Problem (8) have been taken into account.[1]

B. Optimization of p

The problem to be solved is stated as

$$\max_{p} \ F(p) \tag{26a}$$

$$\text{s.t. } AR(p) \geqslant R_{min} \tag{26b}$$

$$P_{min} \leqslant p \leqslant P_{max} \ , \tag{26c}$$

with

$$\begin{aligned} F(p) &= \phi AR(p) - (1-\phi)APC(p) \\ &= \phi N B \lambda_{MS} \log_2\left(1 + \frac{\alpha p}{\beta p + \gamma}\right) - (1-\phi)(\delta p + \zeta) \ , \end{aligned} \tag{27}$$

and

$$\alpha = (1-\epsilon^2)\lambda^2(M - \lambda_{MS}/\lambda) \tag{28}$$

$$\beta = \left[\frac{2}{\eta - 2} - \epsilon^2\right]\lambda^2\lambda_{MS} + \epsilon^2 M \lambda^2 \tag{29}$$

$$\gamma = \Gamma\left(1 + \frac{\eta}{2}\right)\omega\sigma^2/\pi \ , \ \delta = \lambda_{MS} N \mu \frac{T - T_{opt}}{T} \tag{30}$$

$$\begin{aligned} \zeta = &\frac{\lambda N P_{FIX}}{\beta_0} + \frac{\lambda N P_{SYN}}{\beta_1} + \frac{\lambda N \mathcal{P}_0 M}{\beta_1} \\ &+ N\mathcal{P}_1 M \lambda_{MS} + \mathcal{P}_2 \lambda_{MS} + \frac{\lambda T_{opt}}{T} P_{opt} \end{aligned} \tag{31}$$

Studying the first-order derivative of F, we obtain that

$$\frac{\partial F}{\partial p} \geqslant 0 \longleftrightarrow (\alpha+\beta)\beta p^2 + (\alpha + 2\beta)\gamma p + \gamma^2 \leqslant \psi \ , \tag{32}$$

with $\psi = \frac{\phi N B \lambda_{MS} \ln(2)\alpha\gamma}{(1-\phi)\delta}$.

The inequality above is fulfilled for $\underline{p} \leqslant p \leqslant \bar{p}$, with $\underline{p}$ and $\bar{p}$ being the solutions of the associated second order equation. Elaborating, it readily follows that $\underline{p}$ is negative, and thus not feasible, while $\bar{p}$ can be computed as

$$\bar{p} = \frac{-(\alpha+2\beta)\gamma + \sqrt{\gamma^2\alpha^2 + 4\beta(\alpha+\beta)\psi}}{2\beta(\alpha+\beta)} \ . \tag{33}$$

Studying the sign of (33) reveals that $\bar{p}$ is non-negative when

$$\begin{aligned} &\gamma^2\alpha^2 + 4\beta(\alpha+\beta)\psi \geqslant (\alpha+2\beta)^2\gamma^2 \\ &\longleftrightarrow \psi \geqslant \gamma^2 \\ &\longleftrightarrow \frac{\psi N B \ln(2)\alpha}{(1-\phi)\delta\gamma} \geqslant 1 \end{aligned} \tag{34}$$

Thus, when (34) holds, the optimal solution of (26) is given by

$$p^* = \max\{p_0, \min\{\bar{p}, P_{max}\}\} \ , \tag{35}$$

wherein $p_0 = \max\{P_{min}, P_R\}$, with P_R being the power level that fulfills the minimum rate constraint with equality. Elaborating leads to

$$P_R = \frac{\gamma(2^{R_{min}/NB\lambda_{MS}} - 1)}{\alpha - \beta(2^{R_{min}/NB\lambda_{MS}} - 1)} \ . \tag{36}$$

If instead (34) does not hold, then the objective function F is monotonically decreasing for positive p, and thus the solution of (26) is given by $p^* = p_0$.

C. Optimization of M

The problem to be solved is stated as

$$\max_{M} \ F(M) \tag{37a}$$

$$\text{s.t. } AR(M) \geqslant R_{min} \tag{37b}$$

$$M_{min} \leqslant M \leqslant M_{max} \ , \tag{37c}$$

[1] Note that, depending on the values of the parameters a, ψ, γ, δ, and ϕ, it is still possible that the optimal number of component carriers is the maximum or minimum feasible value of N.

with

$$F(M) = \phi AR(M) - (1-\phi)APC(M)$$
$$= \phi NB\lambda_{MS}\log_2\left(1+\frac{\alpha M-\xi}{\beta M+\gamma}\right)-(1-\phi)(\delta M+\zeta), \tag{38}$$

wherein

$$\alpha = (1-\epsilon^2)\ ,\ \beta = \epsilon^2\ ,\ \xi = (1-\epsilon^2)\lambda_{MS}/\lambda \tag{39}$$
$$\delta = \lambda N\mathcal{P}_0/\beta_1 + N\mathcal{P}_1\lambda_{MS} \tag{40}$$
$$\gamma = \Gamma\left(1+\frac{\eta}{2}\right)\omega\sigma^2/(\pi p\lambda^2) + \frac{2p\lambda_{MS}\lambda}{\eta-2} - \epsilon^2\lambda_{MS}/\lambda\ , \tag{41}$$
$$\zeta = \lambda_{MS}N\mu p\frac{T-T_{opt}}{T} + \frac{\lambda NP_{FIX}}{\beta_0} + \frac{\lambda NP_{SYN}}{\beta_1} + \mathcal{P}_2\lambda_{MS} + \frac{\lambda T_{opt}}{T}P_{opt} \tag{42}$$

In order to determine the optimal integer M, we proceed by first relaxing M to a continuous variable. This allows us to study the first-order derivative of F, which, by exploiting the fact that $\alpha+\beta=1$, yields

$$\frac{\partial F}{\partial M} \geqslant 0 \longleftrightarrow \beta M^2 + ((\beta+1)\gamma - \xi\beta)\,M + (\gamma-\xi)\gamma \leqslant \psi\ , \tag{43}$$

with $\psi = \frac{\phi NB\lambda_{MS}\ln(2)(\alpha\gamma+\beta\xi)}{(1-\phi)\delta}$. By a similar reasoning as in the case of the optimization of p, we obtain that the unique feasible solution is given by

$$\bar{M} = \frac{\xi\beta-(\beta+1)\gamma+\sqrt{(\gamma(1-\beta)+\xi\beta)^2+4\beta\psi}}{2\beta}\ , \tag{44}$$

which can be shown to be non-negative whenever $\psi \geqslant \gamma(\gamma-\xi)$. In this case, the integer solution of (37) is given by

$$M^* = \max\{M_0, \min\{\lceil\bar{M}\rceil, M_{max}\}\}\ , \tag{45}$$

wherein $M_0 = \max\{M_{min}, M_R\}$, with M_R being the smallest integer M that fulfills the minimum rate constraint with equality. Elaborating, we obtain

$$M_R = \left\lceil \frac{\xi+\gamma(2^{R_{min}/NB\lambda_{MS}}-1)}{\alpha-\beta(2^{R_{min}/NB\lambda_{MS}}-1)} \right\rceil\ . \tag{46}$$

Instead, if $\psi < \gamma(\gamma-\xi)$, then the objective function F is monotonically decreasing for positive M, and thus the solution of (37) is given by $M^* = M_0$.

Remark 2: The result above shows that, unlike what happens for the number of component carriers, in general the optimum number of antennas to trade-off rate for power consumption is an intermediate value between the maximum and minimum number of antennas.

D. Optimization of λ

The problem to be solved is stated as

$$\max_{\lambda}\ F(\lambda) \tag{47a}$$
$$\text{s.t. } AR(\lambda) \geqslant R_{min} \tag{47b}$$
$$\lambda_{min} \leqslant \lambda \leqslant \lambda_{max}\ . \tag{47c}$$

with

$$F(\lambda) = \phi AR(\lambda) - (1-\phi)APC(\lambda)$$
$$= \phi NB\lambda_{MS}\log_2\left(1+\frac{\alpha\lambda^2-\beta\lambda}{\gamma\lambda^2+\delta\lambda+\upsilon}\right)-(1-\phi)(\nu\lambda+\psi), \tag{48}$$

wherein

$$\alpha = p(1-\epsilon^2)M\ ,\ \beta = p(1-\epsilon^2)\lambda_{MS}\ ,\ \gamma = p\epsilon^2 M \tag{49}$$
$$\delta = \left[\frac{2}{\eta-2}-\epsilon^2\right]p\lambda_{MS}\ ,\ \upsilon = \Gamma\left(1+\frac{\eta}{2}\right)\omega\sigma^2/\pi \tag{50}$$
$$\nu = \frac{N}{\beta_0}P_{FIX} + \frac{N}{\beta_1}P_{SYN} + \frac{N}{\beta_1}\mathcal{P}_0 M + \frac{T_{opt}}{T}P_{opt} \tag{51}$$
$$\psi = N\mu p\lambda_{MS}\frac{T-T_{opt}}{T} + \lambda_{MS}(N\mathcal{P}_1 M + \mathcal{P}_2) \tag{52}$$

Unfortunately, in this case, the study of the first-order derivative of F leads to an equation that can only be solved numerically. Then, let us denote by $\mathcal{L}$ the set containing the feasible stationary points of $F(\lambda)$, which can be computed numerically. Then, the optimal λ can be found by selecting the element of $\mathcal{L}$ that yields the largest value of the objective, namely

$$\lambda^* = \arg\max_{\mathcal{L}\bigcup\{\lambda_{max},\lambda_0\}} F(\lambda)\ , \tag{53}$$

with $\lambda_0 = \max\{\lambda_{min}, \lambda_R\}$, and λ_R equal to the value of λ, which solves the rate constraint with equality. The value λ_R can be determined in closed-form. Indeed, defining $\rho = 2^{R_{min}/NB\lambda_{MS}}-1$, λ_R can be found by solving the equation

$$\lambda_{MS}NB\log_2\left(1+\frac{\alpha\lambda^2-\beta\lambda}{\gamma\lambda^2+\delta\lambda+\upsilon}\right) = R_{min}$$
$$\longleftrightarrow (\alpha-\rho\gamma)\lambda^2 - (\beta-\rho\delta)\lambda - \rho\upsilon = 0\ , \tag{54}$$

which can be seen to admit only one non-negative solution, which is given by

$$\lambda_R = \frac{(\beta+\delta\rho)+\sqrt{(\beta+\delta\rho)^2+4\upsilon\rho(\alpha-\rho\delta)}}{2(\alpha-\rho\delta)}\ . \tag{55}$$

Finally, we can state the overall alternating optimization algorithm as in Algorithm 1, for which the following convergence result holds.

Proposition 2: Algorithm 1 monotonically improves the value of the objective function and converges in the value of the objective.

Proof: In each iteration of Algorithm 1, each optimization variable is updated with the optimal solution of the corresponding subproblem. Thus, the sequence $\{F_n\}$ of the values of the objective is monotonically increasing. Thus, since the objective F is an upper-bounded function, the sequence $\{F_n\}$ must eventually converge. ∎

Running Algorithm 1 for all $\phi \in [0,1]$ provides the Pareto-boundary of Problem (6).

IV. Area rate versus EE trade-off

Following a similar approach as in the previous sections, we tackle Problem (7) by alternating optimization of the variables p, M, λ, and N.

Algorithm 1 Rate versus Power trade-off optimization

Initialize M, p, λ, N to feasible values;
$n = 0$; $F_n = 0$; $F_{-1} = -1$; $\epsilon > 0$; $\phi \in [0, 1]$;
while $F_n - F_{n-1} > \epsilon$ **do**
  Compute a, δ, ψ, γ, ζ in (10), (11), (12);
  if $\phi a \geqslant (1-\phi)(\delta+\psi+\gamma$ **then**
    $N = N_{max}$;
  else if $\phi a < (1-\phi)\delta$ **then**
    $N = \max\left\{N_{min}, \left\lceil \frac{R_{min}}{a} \right\rceil\right\}$;
  else
    Set $\bar{N}$ as the solution of (15);
    $N^* = \max\left\{\max\left\{N_{min}, \left\lceil \frac{R_{min}}{a}\right\rceil\right\}, \min\{\lceil \bar{N} \rceil, N_{max}\}\right\}$;
  end if
  Compute α, β, γ, ζ in (28), (29), (30), (31);
  $\psi = \frac{\phi N B \ln(2) \alpha\gamma}{(1-\phi)\delta}$;
  Compute $\bar{p}$ in (33) and P_R in (36);
  if $\psi \geqslant \gamma^2$ **then**
    $p^* = \max\{\max\{P_{min}, P_R\}, \min\{\bar{p}, P_{max}\}\}$;
  else
    $p^* = \max\{P_{min}, P_R\}$;
  end if
  Compute α, β, ξ, δ, γ, ζ in (39), (40), (41), (42);
  $\psi = \frac{\phi N B \ln(2)(\alpha\gamma+\beta\xi)}{(1-\phi)\delta}$;
  Compute $\bar{M}$ in (44) and M_R (46);
  if $\psi \geqslant \gamma(\gamma - \xi)$ **then**
    $M^* = \max\{\max\{M_{min}, M_R\}, \min\{\lceil \bar{M} \rceil, M_{max}\}\}$;
  else
    $M^* = \max\{M_{min}, M_R\}$;
  end if
  Compute α, β, γ, δ, υ, ν, ψ in (49), (50), (51), (52)
  $\lambda_0 = \max\{\lambda_{min}, \lambda_R\}$, with λ_R in (55);
  $\lambda^* = \arg\max_{\mathcal{L}\bigcup\{\lambda_{max}, \lambda_0\}} F(\lambda)$;
  $F_{n-1} = F_n$;
  $F_n = \phi R(p^*, N^*, M^*, \lambda^*) - (1-\phi) APC(p^*, N^*, M^*, \lambda^*)$;
end while

A. Optimization of N

The problem to be solved is stated as

$$\max_{N} \; F(N) \tag{56a}$$
$$\text{s.t. } AR(N) \geqslant R_{min} \tag{56b}$$
$$N_{min} \leqslant N \leqslant N_{max} \tag{56c}$$

with

$$F(N) = \phi AR(N) + (1-\phi)\text{EE}(N) = Na\left(\phi + \frac{1-\phi}{\delta N + \psi g_0(N) + \gamma g_1(N) + \zeta}\right) , \tag{57}$$

and a given in (10), while

$$\delta = \lambda_{MS}\mu p + \frac{T}{T-T_{opt}}\mathcal{P}_1 M \lambda_{MS}, \psi = \frac{T}{T-T_{opt}} P_{FIX}\lambda \tag{58}$$

$$\zeta = \frac{T\mathcal{P}_2\lambda_{MS} + \lambda T_{opt} P_{opt}}{T - T_{opt}}, \gamma = \frac{T\lambda(P_{SYN} + \mathcal{P}_0 M)}{T - T_{opt}}. \tag{59}$$

Proposition 3: For any functions β_0 and β_1, that are increasing with values in $[1, N]$, the optimal solution of Problem (56) is $N^* = N_{max}$.

Proof: Relaxing N to the continuous domain, we obtain that

$$\frac{\partial F}{\partial N} = a\phi + a(1-\phi)\frac{\psi(g_0(N) - Ng_0^{'}(N)) + \gamma(g_1(N) - Ng_1^{'}(N)) + \zeta}{(\delta N + \zeta)^2}. \tag{60}$$

Moreover, it holds that

$$g_0(N) - Ng_0^{'}(N) = \frac{N}{\beta_0(N)} - \frac{N}{\beta_0^2(N)}\left(\beta_0(N) - N\beta_0^{'}(N)\right) = \frac{N\beta_0^{'}(N)}{\beta_0^2(N)} \geqslant 0 \tag{61}$$

$$g_1(N) - Ng_1^{'}(N) = \frac{N}{\beta_1(N)} - \frac{N}{\beta_1^2(N)}\left(\beta_1(N) - N^2\beta_1^{'}(N)\right) = \frac{N^2\beta_1^{'}(N)}{\beta_1^2(N)} \geqslant 0 , \tag{62}$$

where the last inequalities follow because β_0 and β_1 are increasing functions with N. Thus, (60) is non-negative for any positive value of N, which proves that the objective function F is always monotonically increasing in the feasible set. Hence, the optimal integer N^* is $N^* = N_{max}$. ∎

Remark 3: The result above shows that, unlike what happens for the rate versus power consumption trade-off, when the power consumption metric is replaced by the EE, it is always optimal to transmit employing all available component carriers, regardless of the choice of the functions β_0 and β_1. Thus, employing all available component carriers is the optimal choice for inter-carrier, intra-carrier contiguous, and intra-carrier non-contiguous carrier aggregation. Intuitively, this can be explained noticing that the rate and the EE are not totally contrasting objectives, like the rate and the power consumption. Indeed, the numerator of the EE is equal to the rate. Therefore, even when the priority is completely in favor of the EE, i.e. $\phi = 0$, there is still an incentive to obtain a high rate, and thus use all available component carriers.

B. Optimization of p

The problem to be solved is stated as

$$\max_{p} \; F(p) \tag{63a}$$
$$\text{s.t. } AR(p) \geqslant R_{min} \; , \; P_{min} \leqslant p \leqslant P_{max} \; , \tag{63b}$$

with

$$F(p) = \phi AR(p) + (1-\phi)\text{EE}(p) = NB\lambda_{MS}\log_2\left(1 + \frac{\alpha p}{\beta p + \gamma}\right)\left(\phi + \frac{(1-\phi)}{\delta p + \zeta}\right), \tag{64}$$

and α, β, γ given in (28), (29), (30), while

$$\delta = \lambda_{MS} N\mu \tag{65}$$

$$\zeta = \frac{\lambda N P_{FIX}}{\beta_0(N)} + \frac{\lambda N(P_{SYN} + \mathcal{P}_0 M)}{\beta_1(N)} + N\mathcal{P}_1 M\lambda_{MS} + \mathcal{P}_2\lambda_{MS} \tag{66}$$

The stationarity condition for F is given by

$$\frac{\partial F}{\partial p}=0\longleftrightarrow\frac{\ln(2)\alpha\gamma}{((\alpha+\beta)p+\gamma)(\beta p+\gamma)}\left(\phi+\frac{1-\phi}{\delta p+\zeta}\right)=\frac{(1-\phi)\delta}{(\delta p+\zeta)^2}\log_2\left(1+\frac{\alpha p}{\beta p+\gamma}\right)\tag{67}$$

Unfortunately, (67) does not admit a closed-form solution. On the other hand, it is possible, however, to guarantee that (67) admits a unique solution, which coincides with the global maximizer of the function F. To this end, let us first provide the following lemma.

Lemma 1: Let $g:\mathbb{R}_0^+\rightarrow\mathbb{R}_0^+$ be a differentiable, increasing, and strictly S-shaped function,[2] with $f(0)=0$. Then, the ratio $g(x)=\frac{f(x)}{ax+b}$, with a,b positive coefficients, is a strictly pseudo-concave function.

Proof: Denote by x^* the point below which f is strictly convex, and above which it is strictly concave. We will prove the result by separately showing that g is strictly pseudo-concave for $x\geqslant x^*$ and strictly increasing for $x<x^*$. This clearly implies strict pseudo-concavity of g for all $x\geqslant 0$, since g is also differentiable in x^*. For $x\geqslant x^*$, g is the ratio of a strictly concave function over a linear function, and therefore it is strictly pseudo-concave. Next, we have to show that for $x<x^*$, g is strictly increasing. To see this let us compute the first derivative of g, which yields

$$\frac{dg}{dx}=\frac{(ax+b)f'(x)-af(x)}{(ax+b)^2}=\frac{a(xf'(x)-f(x))+f'(x)b}{(ax+b)^2}\ .\tag{68}$$

Now, for $x<x^*$, we have that f is strictly convex, thus implying that $(y-x)\,f'(x)<f(y)-f(x)$, for all $x,y<x^*$. Setting $y=0$, we obtain the condition $xf'(x)>f(x)$, for all $x<x^*$, which shows that (68) is strictly positive. Then, we have that, for $x<x^*$, $g(x)$ is strictly increasing, and hence, the proof. ■

Equipped with Lemma 1, we can prove the following result.

Proposition 4: Define $d=(\phi\zeta+1-\phi)/\delta\phi$, and assume $2\alpha d\beta+2d\beta^2-\alpha\gamma-2\beta\gamma>0$. Then, the stationarity condition in (67) admits a unique solution, which is the global maximizer of the function F in (64).

Proof: Let us rewrite the objective function (64) as

$$F(p)=NB\lambda_{MS}\delta\phi\frac{(p+d)\log_2(f(p))}{\delta p+\zeta}\ ,\tag{69}$$

with $f(p)=1+\frac{\alpha p}{\beta p+\gamma}$. Now, let us consider only the numerator of F, which we have defined as the function $g(p)=NB\lambda_{MS}\delta\phi(p+d)\log_2(f(p))$. The next step of the proof is to show that g is an S-shaped function, i.e. it has a unique inflection point p_0, being convex for $p\leqslant p_0$ and concave for $p\geqslant 0$. Studying the sign of the second derivative of g, after some elaborations, we obtain the condition

$$\frac{\partial^2 g}{\partial p^2}\geqslant 0\longleftrightarrow 2f^{'}(p)f(p)\geqslant\left[(f^{'}(p))^2-f^{''}(p)f(p)\right](x+d)\tag{70}$$

Plugging in the first and second derivative of f, and elaborating yields that

$$\begin{aligned}2(\alpha+\beta)p+2\gamma&\geqslant(x+d)\left[\frac{\alpha\gamma+2\beta(\alpha+\beta)p+2\beta\gamma}{\beta p+\gamma}\right]\\2\gamma^2-d\alpha\gamma-2\beta\gamma d&\geqslant p(2\alpha\beta d+2d\beta^2-\alpha\gamma-2\beta\gamma)\\p&\leqslant\frac{\gamma(2\gamma-d\alpha-2\beta d)}{2\alpha d\beta+2d\beta^2-\alpha\gamma-2\beta\gamma}=p_0\ ,\end{aligned}\tag{71}$$

wherein in the last step we have exploited the assumption that $2\alpha d\beta+2d\beta^2-\alpha\gamma-2\beta\gamma\geqslant 0$. The function g is thus strictly convex for $p\leqslant p_0$ and strictly concave for $p\geqslant p_0$, i.e. it is S-shaped. As a consequence, by virtue of Lemma 1, it holds that the objective F is strictly pseudo-concave, since it is the ratio between the S-shaped function g, which fulfills all assumptions of Lemma 1 over the problem feasible set, and the affine function $\delta p+\zeta$. Thus, the result follows because a strictly pseudo-concave function admits a unique global maximizer, which coincides with its unique stationary point. ■

Remark 4: It should be noted that there was no need to make any assumption on the sign of $\gamma(2\gamma-d\alpha-2\beta d)$. Indeed, if $\gamma(2\gamma-d\alpha-2\beta d)$, then $p_0<0$, and the function g will be strictly concave over the problem feasible set, which shows again the strict pseudo-concavity of the objective F.[3]

Remark 5: As for the practicality of the assumption $2\alpha d\beta+2d\beta^2-\alpha\gamma-2\beta\gamma>0$, we observe the following. The coefficient γ is proportional to the noise power σ^2, while the coefficients β and d tend to assume large values for practical choices of the system parameters. Indeed, β is proportional to the number of antennas, base station density, and MS density, while d is proportional to the static power consumption ζ, which is in turn proportional to the number of component carriers, antennas, and MS density. Thus, it is practical to assume that $2d\beta>\gamma$ and $d\beta>\gamma$, which together imply the assumption $2\alpha d\beta+2d\beta^2-\alpha\gamma-2\beta\gamma\geqslant 0$.

Based on the results above, denoting by $\bar{p}$ the unique solution p^* of (67), the unique solution of Problem (63) is obtained by taking into account the constraints of the problem, which leads to

$$p^*=\max\{p_0,\min\{\bar{p},P_{max}\}\}\ ,\tag{72}$$

with $p_0=\max\{P_{min},P_R\}$, where P_R is the power level that fulfills the minimum rate constraint with equality, which can be found as

$$P_R=\frac{\gamma(2^{R_{min}/NB\lambda_{MS}}-1)}{\alpha-\beta(2^{R_{min}/NB\lambda_{MS}}-1)}.\tag{73}$$

C. Optimization of M

The problem to be solved is stated as

$$\max_{p}\ F(M)\tag{74a}$$

$$\text{s.t. } AR(M)\geqslant R_{min}\tag{74b}$$

$$M_{min}\leqslant M\leqslant M_{max}\ ,\tag{74c}$$

[2] An increasing, positive function is said to be strictly S-shaped, if there exists a point below which it is strictly convex, and a point above which it is strictly concave.

[3] Recall that the ratio between a strictly concave function and an affine function is strictly pseudo-concave [6].

with

$$F(M) = \phi AR(M) + (1-\phi)\mathrm{EE}(M)$$
$$= NB\lambda_{MS}\log_2\left(1+\frac{\alpha M-\xi}{\beta M+\gamma}\right)\left(\phi+\frac{(1-\phi)}{\delta M+\zeta}\right), \quad (75)$$

and α, β, ξ, γ given in (39), (41), and (42), respectively, while

$$\delta = \lambda N\mathcal{P}_0/\beta_1(N) + N\mathcal{P}_1\lambda_{MS} \quad (76)$$
$$\zeta = \lambda_{MS}N\mu p + \frac{\lambda NP_{FIX}}{\beta_0(N)} + \frac{\lambda NP_{SYN}}{\beta_1(N)} + \mathcal{P}_2\lambda_{MS} \quad (77)$$

Let us observe that the rate is an increasing function of the number of antennas M, only if $\frac{\alpha M-\xi}{\beta M+\gamma}$ is increasing with M, which leads to the condition $q = \alpha\gamma + \beta\xi > 0$. In the following, we focus on this scenario, which is that of most practical relevance, since the opposite case $q < 0$ would trivially yield that the optimal number of antennas is $M = 1$. Moreover, it should be noted that the only negative term in q is the coefficient $-\epsilon^2\lambda_{MS}/\lambda$, which appears in γ. and that is proportional to the hardware impairment factor ϵ^2, which is close to 0 for practical hardware. Thus, we can assume that $q > 0$.

In order to analyze the objective F, let us temporarily relax M to a continuous variable. Then, it is possible to consider the derivative of (75) and set it to 0, which yields

$$\frac{\partial F}{\partial M} = 0 \longleftrightarrow \frac{\ln(2)\alpha\gamma}{((\alpha+\beta)M+\gamma-\xi)(\beta M+\gamma)}\left(\phi+\frac{1-\phi}{\delta M+\zeta}\right)$$
$$= \frac{(1-\phi)\delta}{(\delta M+\zeta)^2}\log_2\left(1+\frac{\alpha M-\xi}{\beta M+\gamma}\right) \quad (78)$$

Similarly to (67), a closed-form solution of (78) is not available. Nevertheless, it is possible to ensure that (78) has a unique solution, which coincides with the global maximizer of (75). Specifically, the following proposition holds.

Proposition 5: Define $d = (\phi\zeta + 1 - \phi)/\delta\phi$, and assume $q + 2\beta\alpha d - 2\alpha\gamma > 0$. Then, the stationarity condition in (78) admits a unique solution, which is the global maximizer of the function F in (75).

Proof: Let us rewrite the function in (75) as

$$F(M) = NB\delta\phi\frac{(M+d)\log_2(f(M))}{\delta M+\zeta}, \quad (79)$$

with $f(M) = 1 + \frac{\alpha M-\xi}{\beta M+\gamma}$. Now, let us prove that the numerator of F, which we have defined as the function $g(M) = NB\lambda_{MS}\delta\phi(M+d)\log_2(f(M))$, is an S-shaped function. Studying the sign of the second derivative of g, after similar steps as in the proof of Proposition 4, we obtain the condition

$$2q(aM-\xi)(\beta M+\gamma) \geqslant (M+d)(q^2 + 2\beta\alpha(\alpha M-\xi))$$
$$2\beta qd\xi - 2q\xi\gamma - qd \geqslant Mq(q + 2d\beta\alpha - 2\alpha\gamma)$$
$$M \leqslant \frac{2\beta qd\xi - 2q\xi\gamma - qd}{q(q+2d\beta\alpha-2\alpha\gamma)} = M_0\ . \quad (80)$$

where, in the last step, we have exploited the assumption that $q > 0$ and $(q+2d\beta\alpha-2\alpha\gamma) > 0$. The function g is thus strictly convex for $M \leqslant M_0$ and strictly concave for $M \geqslant M_0$, i.e. it is S-shaped. As a consequence, applying again Lemma 1, it holds that the function F is strictly pseudo-concave, and the result follows. ■

Remark 6: As for the practicality of the assumption $(q + 2d\beta\alpha - 2\alpha\gamma) > 0$, the condition can be rewritten as

$$2\epsilon^2\frac{\lambda_{MS}}{\lambda} + 2d\epsilon^2 > 2\frac{\lambda_{MS}}{(\eta-2)\lambda} + \Gamma\left(1+\frac{\eta}{2}\right)\frac{\omega\sigma^2}{\pi p\lambda^2}, \quad (81)$$

which approximately reduces to

$$d\epsilon^2 > \frac{\lambda_{MS}}{(\eta-2)\lambda}, \quad (82)$$

because the term that is proportional to the noise power σ^2 is negligible compared to the other term on the right-hand-side, while $\epsilon^2\lambda_{MS}/\lambda$ is negligible compared to the coefficient d, which is proportional to ζ, i.e., the power consumption of the base station that does not depend on the number of antennas. To elaborate further, we argue that, despite being multiplied by ϵ^2, the coefficient d is still likely to be larger than the right-hand-side of (82). To see this, let us consider that a typical value for the coefficient P_{FIX} that measures the power consumption of base-band signal processing is of the order of 300 W.[4] Then, a system employing even a relatively low number of component carriers, e.g. $N = 10$, with a base station density of $\lambda = 10$, would consume $P_{FIX}N\lambda = 3 * 10^4$ W in terms of static power consumption. On the other hand, even at peak hours with a MS density of the order of $\lambda_{MS} = 200$, in the same system, we would have $\lambda_{MS}/\lambda = 20$. Thus, recalling that η is a number larger than 2, but of the order of units, and that typical hardware of good quality has a factor ϵ^2 of the order of 10^{-2}, we see that (82) holds by a large margin.

Equipped with the above results, let us denote by $\bar{M}$ the unique solution of (78). Then, the optimal integer solution M^* of Problem (74) is found by mapping back into the integer domain and accounting for the constraints, which leads to

$$M^* = \max\{M_0, \min\{\lceil\bar{M}\rceil, M_{max}\}\}, \quad (83)$$

with $M_0 = \max\{M_{min}, M_R\}$, and M_R being the smallest integer M that fulfills the minimum rate constraint with equality, i.e.

$$M_R = \left\lceil\frac{\xi + \gamma(2^{R_{min}/NB\lambda_{MS}} - 1)}{\alpha - \beta(2^{R_{min}/NB\lambda_{MS}} - 1)}\right\rceil \quad (84)$$

D. Optimization of λ

The problem to be solved is stated as

$$\max_{\lambda}\ F(\lambda) \quad (85a)$$
$$\text{s.t. } AR(\lambda) \geqslant R_{min} \quad (85b)$$
$$\lambda_{min} \leqslant \lambda \leqslant \lambda_{max}\ . \quad (85c)$$

with

$$F(\lambda) = \phi AR(\lambda) - (1-\phi)\mathrm{EE}(\lambda)$$
$$= NB\lambda_{MS}\log_2\left(1+\frac{\alpha\lambda^2-\beta\lambda}{\gamma\lambda^2+\delta\lambda+\upsilon}\right)\left(\phi+\frac{1-\phi}{\nu\lambda+\psi}\right), \quad (86)$$

[4]This is the value that will be used in the numerical simulations.

and α, β, γ, δ, υ given in (49), (50), while

$$\nu = \frac{N}{\beta_0(N)} P_{FIX} + \frac{N}{\beta_1(N)} P_{SYN} + \frac{N}{\beta_1(N)} \mathcal{P}_0 M \quad (87)$$

$$\psi = N\mu p \lambda_{MS} + \lambda_{MS}(N\mathcal{P}_1 M + \mathcal{P}_2) \quad (88)$$

Unfortunately, setting the first-order derivative of F with respect to λ to 0 yields an equation that can only be solved numerically, and for which it is not clear if only 1 solution exists. Nevertheless, proceeding similarly to the rate-power consumption trade-off case, let us denote by $\mathcal{L}$ the set containing the feasible stationary points of $F(\lambda)$, which can be computed numerically. Then, the optimal λ can be found by selecting the element of $\mathcal{L}$ that yields the largest value of the objective, namely

$$\lambda^* = \arg \max_{\mathcal{L} \bigcup \{\lambda_{max}, \lambda_0\}} F(\lambda) \ , \quad (89)$$

with $\lambda_0 = \max\{\lambda_{min}, \lambda_R\}$ and λ_R equal to the value of λ that solves the rate constraint with equality. Following similar steps as for the rate-power trade-off, the value λ_R can be computed as in (55).

Finally, we can state the overall alternating optimization algorithm as in Algorithm 2, for which the following convergence result holds.

Proposition 6: Algorithm 2 monotonically improves the value of the objective function and converges in the value of the objective.

Proof: The proof follows upon the same lines as for Algorithm 1. ■

Running Algorithm 2 for all $\phi \in [0,1]$ provides the Pareto-boundary of Problem (7).

Algorithm 2 Rate versus EE trade-off optimization

Initialize M, p, λ, N to feasible values;
$n = 0$; $F_n = 0$; $F_{-1} = -1$; $\epsilon > 0$; $\phi \in [0,1]$;
$N = N_{max}$;
while $F_n - F_{n-1} > \epsilon$ **do**
  Compute α, β, γ in (28), (29), (30) and δ, ζ in (65), (66);
  Compute $\bar{p}$ as the unique solution of (67);
  Compute P_R as in (73);
  $p^* = \max\{\max\{P_{min}, P_R\}, \min\{\bar{p}, P_{max}\}\}$;
  Compute α, β, ξ, γ in (39), (41), and δ, ζ in (76), (77);
  Compute $\bar{M}$ as the unique solution of (78);
  Compute M_R as in (84);
  $M^* = \max\{\max\{M_{min}, M_R\}, \min\{\lceil \bar{M} \rceil, M_{max}\}\}$;
  Compute α, β, γ, δ, υ, in (49), (50), (51), (52) and ν, ψ in (87), (88)
  $\lambda_0 = \max\{\lambda_{min}, \lambda_R\}$, with λ_R in (55);
  $\lambda^* = \arg\max_{\mathcal{L} \bigcup \{\lambda_{max}, \lambda_0\}} F(\lambda)$;
  $F_{n-1} = F_n$;
  $F_n = \phi R(p^*, N^*, M^*, \lambda^*) - (1 - \phi) APC(p^*, N^*, M^*, \lambda^*)$;
end while

V. Numerical Results

The performance of the developed resource allocation methods has been analyzed by numerical simulation in a multi-cell multi-carrier, massive MIMO network with the following parameters:

- $B = 20\,\text{MHz}$, $\sigma^2 = 6.32 * 10^{-13}$, $\eta = 4$, $\omega = 131.4\,\text{dB}$
- $\epsilon = 0.05$, $T = 2 * 10^{-5} s$, λ_{MS}=25 /km^2
- $P_{FIX} = 300\,\text{W}$, $P_{SYN} = 34\,\text{W}$, $\mathcal{P}_0 = 4.49\,\text{W}$
- $f_c = 4\,\text{GHz}$, $\eta_f = 12.8\,\text{Gflop/W}$, $\tau = 1/B$
- $\mathcal{A} = 3.5 * 10^{-8}$, $R_{BS} = 100\,\text{Mb/s}$, $R_{min} = 100\,\text{Mb/s}$
- $N_{min} = 1$, $N_{max} = 64$, $M_{min} = 64$, $M_{max} = 512$
- $\lambda_{min} = 10/\text{km}^2$, $\lambda_{max} = 100/\text{km}^2$, $P_{min} = -20\,\text{dBW}$, $P_{max} = 20\,\text{dBW}$

In all numerical results, the displayed rate function is the area rate in (1) normalized to the mobile stations density λ_{MS}, which is a constant for all considered problems and thus does not affect the optimal resource allocations. Moreover, we have set $\beta_0(N) = \beta_1(N) = 1$, i.e., inter-band carrier aggregation. This represents a worst-case power consumption model, because it leads to a power consumption model that scales linearly with the component carriers, while with the other carrier aggregation models, the power consumption would scale sub-linearly with the component carriers. As for the MS density, at first we set it to $\lambda_{MS} = 25\,\text{MS/km}^2$, which models a lightly-loaded network. In this scenario, Figs. 1 and 2 show the Pareto optimal frontier of Problems (6) and (7), respectively, as obtained by means of Algorithm 1. Moreover, each figure also shows specific points on the Pareto frontier, with the associated optimized resource allocation. In addition to the optimized values of N, p, M, and λ, the value K computed based on (5) is also reported.

Inspecting Fig. 1, one can see that, in the extreme case in which the rate priority is set to 0 ($\phi = 0$), all resources are set to the minimum values that fulfill the rate requirements, because in this case the problem reduces to power consumption minimization subject to minimum rate constraints. On the other hand, when the rate priority approaches 1 ($\phi \to 1$), all resources are allocated to their maximum values, because in this case there is no concern about power consumption, and all resources are allocated to provide the best possible rate. Moreover, as we increase the priority of the rate, the first variable increased to support the higher rate is the number of component carriers, which rapidly reaches its maximum value.

Thus, for the considered realistic setup, the system parameters are such that $N = N_{max}$ is the optimal allocation, even if inter-band contiguous carrier aggregation is assumed here, unless a very low rate priority is used, in which case the minimum feasible number of component carriers is employed. After $N = N_{max}$ has been reached, the number of antennas, the base station density, and the transmit power are increased too. The increase of the rate is observed to be more significant when the number of component carriers is increased, then less and less significant as the other resources are increased. At the same time, the increase in power consumption is more limited when the number of component carriers are increased, then more and more significant as the other resources are increased. This shows how increasing the component carriers is the key strategy to increase the rate with a lower power consumption increase than the other resources. These results can be justified by the following considerations:

1) The number of component carriers N is the only variable that affects the rate linearly, while all other resources affect the rate logarithmically. On the other hand, all resources affect the power consumption linearly. As a result, when the priority of the rate in increased, the optimal approach is to increase the number of component carriers, since this yields the largest increase of the rate, while causing a similar power consumption increase than the other variables.
2) The number of antennas M appears both at the numerator and the denominator of the argument of the logarithm in the rate function. However, at the numerator, M multiplies the term $(1-\epsilon)^2$, while at the denominator, it multiplies the term ϵ^2. Since ϵ is typically a small value ($\epsilon^2 = 0.05$ in our simulations), the term at the numerator is much more significant than that at the denominator, and thus it can be claimed that M affects the argument of the logarithm approximately linearly. This explains why the number of antennas M is the second resource to be increased when the priority of the rate is increased.
3) On the other hand, the transmit power p and the base station density λ appear at the denominator of the argument of the logarithm, in a term that is proportional to the inverse of the SNR. Thus, in practical scenarios, in which the system operates at moderate-to-high SNRs, the impact of both p and λ is less significant than that of the component carriers and the number of antennas. Thus, these two resources are the last to be increased as the rate priority increases.

As for the Pareto frontier of the rate versus EE trade-off shown in Fig. 2, similar considerations can be made. However, in this case, as analytically shown, the optimal number of component carriers N is always equal to the maximum value N_{max}. For this reason the initial value of the rate is much higher than the minimum requirement, even for very low rate priorities. Moreover, the increase of the rate as its priority approaches 1 is not as significant as in Fig. 1. due to the fact that $N = N_{max}$ for all points of the frontier. Next, as the rate priority increases, the number of antennas is the resource that is increased first, followed by the base station density and transmit power, as for the rate-power consumption trade-off shown in Fig. 1.

Finally, a heavily-loaded network is considered in Figs. 3 and 4, which is modeled by setting $\lambda_{MS} = 160\,\text{MS/km}^2$. All other parameters, as well as the choice of the functions β_0 and β_1 is the same as in Figs. 1 and 2. Also in this heavily-loaded scenario, a similar trend of the Pareto frontiers as in Figs. 3 and 4 is observed. The only significant difference lies in the fact that the base station density becomes a more relevant variable as a consequence of the larger value of the MS density. Indeed, as we increase the rate priority, it is observed that the base station density increases faster than in the lightly-loaded scenario in order to support the larger number of MSs, which causes a faster increase of the rate as its priority factor approaches 1.

VI. Conclusions

This work has considered a multi-cell, multi-carrier, massive MIMO network in which carrier aggregation is employed. Two trade-off problems have been studied, namely the rate versus power consumption trade-off and the rate versus EE trade-off. Taking a weighted sum utility approach, alternating maximization algorithms have been derived to compute the Pareto frontiers of both problems. For each algorithm, closed-form results have been provided as to the optimal allocation of the individual resources, considering three main classes of carrier aggregation techniques, namely inter-carrier, intra-carrier contiguous, and intra-carrier non-contiguous.

The obtained results show that multi-carrier transmissions are a very effective way of improving the rate and EE of a wireless network, while guaranteeing a limited power consumption. Simulation results show that higher values of rate and EE can be obtained at a lower power consumption if the maximum number of component carriers is increased, compared to the choice of increasing other radio resources like the number of antennas, the base station density, and the transmit power. This behavior has been observed in both lightly-loaded and heavily- loaded networks, and provide useful guidance for radio resource allocation in both scenarios. Moreover, the numerical results have been obtained assuming inter-band carrier aggregation, which is the most power-consuming carrier aggregation method, among the three considered carrier aggregation techniques, although the easiest to be implemented in practice. Even more favorable results in favor of using more component carriers than other resources would be seen for intra-band approaches.

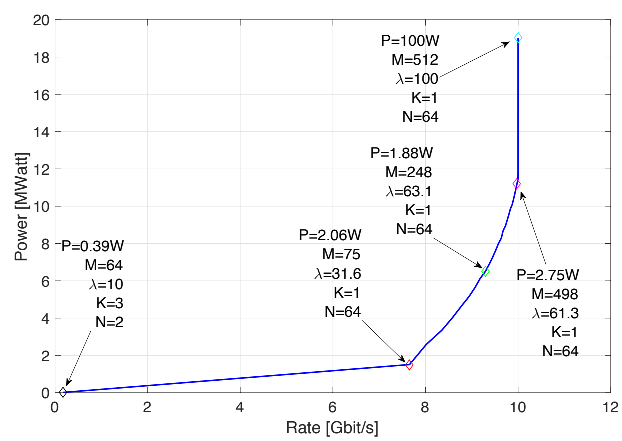

Fig. 1. Rate versus Power consumption. $\lambda_{MS} = 25\,\text{MS/km}^2$.

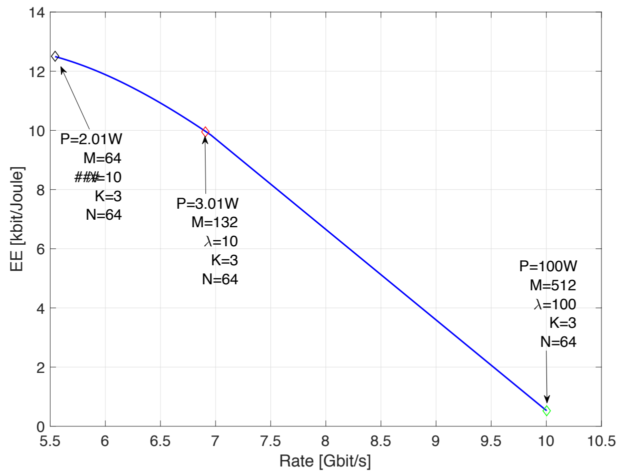

Fig. 2. Rate versus EE. $\lambda_{MS} = 25\,\text{MS/km}^2$.

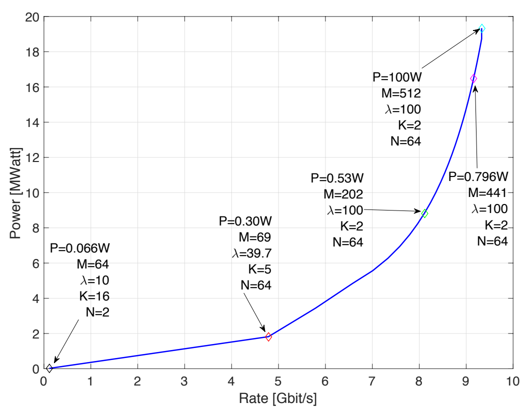

Fig. 3. Rate versus Power consumption. $\lambda_{MS} = 160\,\text{MS/km}^2$.

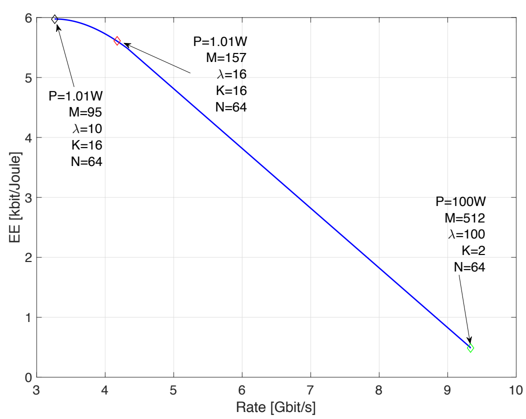

Fig. 4. Rate versus EE. $\lambda_{MS} = 160\,\text{MS/km}^2$.

References

[1] D. López-Pérez, A. De Domenico, N. Piovesan, G. Xinli, H. Bao, S. Qitao, and M. Debbah, "A survey on 5g radio access network energy efficiency: Massive mimo, lean carrier design, sleep modes, and machine learning," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2022.

[2] GSMA, "Energy efficiency, an overview," *https://www.gsma.com/futurenetworks/wiki/energy-efficiency-2/*, 2019.

[3] NGMN Alliance, "5G white paper," *https://ngmn.org/wp-content/uploads/NGMN-5G-White-Paper-2.pdf*, 2020.

[4] Huawei Technologies Co. Ltd., "Green 5G: Building a sustainable world," *Technical Report*, vol. [Online]. Available: https://www.huawei.com/en/public- policy/green- 5g- building- a- sustainable- world, 2020.

[5] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, p. 21, 2015.

[6] A. Zappone and E. A. Jorwsieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185 – 399, 2015.

[7] Y. Hao, Q. Ni, H. Li, and S. Hou, "Energy and spectral efficiency tradeoff with user association and power coordination in massive mimo enabled HetNets," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2091 – 2094, 2016.

[8] Z. Zhang, Z. Chen, M. Shen, and B. Xia, "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 848 – 863, 2016.

[9] W. Xu, J. Liu, S. Jin, and X. Dong, "Spectral and energy efficiency of multi-pair massive MIMO relay network with hybrid processing," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3794 – 3808, 2017.

[10] Y. Hao, Q. Ni, H. Li, and S. Hou, "On the energy and spectral efficiency tradeoff in massive MIMO-enabled hetnets with capacity-constrained backhaul links," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4720 – 4733, 2017.

[11] A. Zappone, E. Björnson, L. Sanguinetti, and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Transactions on Signal Processing*, vol. 65, no. 11, pp. 2844 – 2859, 2017.

[12] M. Arisaka and S. Sugiura, "Energy-versus-bandwidth-efficiency trade-off in spatially modulated massive MIMO downlink," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 197 – 200, 2019.

[13] M. Bashar, H. Q. Ngo, K. Cumanan, A. G. Burr, P. Xiao, E. Björnson, and E. G. Larsson, "Uplink spectral and energy efficiency of cell-free massive MIMO with optimal uniform quantization," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 223 – 245, 2021.

[14] L. You, J. Xiong, A. Zappone, W. Wang, and X. Gao, "Spectral efficiency and energy efficiency tradeoff in massive MIMO downlink transmission with statistical CSIT," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2645 – 2659, 2020.

[15] H. V. Nguyen, V.-D. Nguyen, O. A. Dobre, S. K. Sharma, S. Chatzinotas, B. Ottersten, and O.-S. Shin, "On the spectral and energy efficiencies of full-duplex cell-free massive mimo," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1698 – 1718, 2020.

[16] C. Li, J. Zhang, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2505 – 2517, 2014.

[17] M. D. Renzo, A. Zappone, T. T. Lam, and M. Debbah, "System-level modeling and optimization of the energy efficiency in cellular networks - a stochastic geometry framework," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2539 – 2556, 2018.

[18] ——, "Spectral-energy efficiency pareto front in cellular networks: A stochastic geometry framework," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 424 – 427, 2019.

[19] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059 – 3075, 2015.

[20] E. Björnson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 832 – 847, 2016.

[21] Y. Xin, D. Wang, J. Li, H. Zhu, J. Wang, and X. You, "Area spectral efficiency and area energy efficiency of massive MIMO cellular systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3243 – 3245, 2016.

[22] Z. Liu, W. Du, and D. Sun, "Energy and spectral efficiency tradeoff for massive MIMO systems with transmit antenna selection," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4453 – 4457, 2017.

[23] W. Liu, S. Han, and C. Yang, "Energy efficiency scaling law of massive MIMO systems," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 107 – 121, 2017.

[24] J. Qian, C. Masouros, and M. Matthaiou, "Multi-pair two-way massive MIMO relaying with zero forcing: Energy efficiency and power scaling

laws," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1417 – 1431, 2020.

[25] A. Papazafeiropoulos, H. Q. Ngo, P. Kourtessis, S. Chatzinotas, and J. M. Senior, "Towards optimal energy efficiency in cell-free massive MIMO systems," *IEEE Transactions on Green Communications and Networking*, 2021.

[26] V. Radhakrishnan, O. Taghizadeh, and R. Mathar, "Energy efficient full duplex massive MIMO multi-carrier bidirectional communication with hardware impairments," *23rd International ITG Workshop on Smart Antennas, WSA 2019*, 2019.

[27] V. Radhakrishnan, O. Taghizadeh, and R. Matha, "Rate splitting for massive MIMO multi-carrier system using full duplex decode and forward relay with hardware impairments," *https://arxiv.org/abs/1907.13381*, 2019.

[28] M. Lauridsen, H. Wang, and P. Mogensen, "LTE UE energy saving by applying carrier aggregation in an hetnet scenario," *2017 IEEE 77th Vehicular Technology Conference, VTC Spring 2017*, 2017.

[29] E. Dahlman, S. Parkvall, and J. Skold, "5G NR: The next generation wireless access technology," *Academic Press*, 2018.

[30] F. Liu, K. Zheng, W. Xiang, and H. Zhao, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 197 – 207, 2014.

[31] G. Yu, Q. Chen, R. Yin, H. Zhang, and G. Y. Li, "Joint downlink and uplink resource allocation for energy-efficient carrier aggregation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3207 – 3218, 2015.

[32] Q. Chen, G. Yu, R. Yin, and G. Y. Li, "Energy-efficient user association and resource allocation for multistream carrier aggregation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6366 – 6376, 2016.

[33] E. Chavarria-Reyes, I. F. Akyildiz, and E. Fadel, "Energy-efficient multi-stream carrier aggregation for heterogeneous networks in 5g wireless systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7432 – 7443, 2016.

[34] X. Wang, S. Thota, M. Tornatore, H. H. L. H. S. Chung, S. Park, and B. Mukherjee, "Energy-efficient virtual base station formation in optical-access-enabled cloud ran," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1130 – 1139, 2016.

[35] F. Foukalas and T. A. Tsiftsis, "Energy efficient power allocation for carrier aggregation in heterogeneous networks: Partial feedback and circuit power consumption," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 3, pp. 623 – 634, 2018.

[36] L. Sharma, J.-M. Liang, and S.-L. Wu, "Energy-efficient resource scheduling within DRX cycles for LTE-A networks with carrier aggregation," *IEEE Access*, vol. 6, pp. 28 501 – 28 513, 2018.

[37] D. López-Pérez, A. D. Domenico, N. Piovesan, X. Geng, H. Bao, and M. Debbah, "Energy Efficiency of Multi-Carrier Massive MIMO Networks: Massive MIMO Meets Carrier Aggregation," in *IEEE Globecom 2021*, Madrid, Spain, Dec. 2021.

[38] Y. Hmamouche, M. Benjillali, S. Saoudi, H. Yanikomeroglu, and M. D. Renzo, "New trends in stochastic geometry for wireless networks: A tutorial and survey," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1200–1252, 2021.