Stable Matching for Wireless URLLC in Multi-Cellular, Multi-User Systems

Tom Hößler, *Student Member, IEEE*, Philipp Schulz, *Student Member, IEEE*, Eduard A. Jorswieck, *Fellow, IEEE*, Meryem Simsek, *Senior Member, IEEE*, and Gerhard P. Fettweis, *Fellow, IEEE*

***Abstract*— Ultra-Reliable Low-Latency Communications (URLLC) are considered as one of the key services of the upcoming fifth generation (5G) of wireless communications systems. Enabling URLLC is especially challenging due to the strict requirements in terms of latency and reliability. Multi-connectivity is a powerful approach to increase reliability. However, most of the current research is restricted to single-user scenarios, neglecting the challenges of multi-cellular, multi-user systems, i.e., interference and the competition for limited resources. In this article, we develop analytic comparisons of different connectivity approaches, showing that multi-connectivity may not always be optimal in the considered scenario. Moreover, we propose and evaluate novel resource allocation approaches based on stable matching theory to enable wireless URLLC. We extend the pure many-to-one stable matching procedure by utilizing the optimal connectivity approach for each user, optimizing the maximum number of matched resources, and providing a resource reservation mechanism for users suffering from bad channel conditions. System-level simulations demonstrate that the proposed algorithm outperforms baseline resource allocation approaches in outage probability by up to three orders of magnitude. Even in a highly loaded system, an outage probability in the range of 10^{-5} is achieved.**

***Index Terms*— 5G, multi-connectivity, reliability, wireless systems, stable matching.**

I. Introduction

THE 5G of wireless communications networks is expected to enable a variety of new applications, comprising the key services enhanced Mobile Broadband (eMBB), massive Machine-Type Communications (mMTC), and Ultra-Reliable Low-Latency Communications (URLLC). Among these services, URLLC are especially challenging, because it depends on the simultaneous fulfillment of strict requirements in terms of latency and reliability. This combination is crucial for mission-critical applications such as autonomous driving, wireless factory automation, or the Tactile Internet [1]. In wireless communications, reliability[1] is often interpreted as the success probability of transmitting a packet within a required maximum time. According to the International Telecommunication Union (ITU) and 3rd Generation Partnership Project (3GPP), URLLC services require a reliability of $1 - 10^{-5}$ for delivering a 32-byte packet within 1ms [3], [4]. To achieve high reliability over fading channels, diversity is commonly accepted to be key, mainly classified into space, time, and frequency diversity. To obtain space and/or frequency diversity, the simultaneous connection to multiple wireless links from different base stations and, if possible, at different frequencies is discussed [5]–[7]. This approach is known as multi-connectivity. However, we demonstrated in [8] that adding resources may not necessarily improve the reliability in multi-cellular, multi-user systems due to interference and the competition for limited resources. In this paper, we significantly extend our studies on multi-connectivity in multi-cellular, multi-user systems. First, we provide an analytic comparison of different connectivity approaches, with analysis of connectivity options from a single user's perspective. Here, we focus on a single user's perspective, since URLLC aims to guarantee reliability to each individual user and, in a multi-user system, to as many individual users as possible, if not to all users. Second, we propose a matching theory-based algorithm, that minimizes resource consumption while guaranteeing URLLC service requirements. Lastly, we introduce an extension to the proposed algorithm, which assigns the remaining resources, if the matching theory-based algorithm does not satisfy all users' requirements. The contributions of this article are summarized as follows:

- A multi-cellular, multi-user system with URLLC traffic is studied, i.e., all users have a stringent reliability requirement under a given latency budget.
- Different connectivity approaches, comprising single-connectivity, multi-connectivity, and joint transmission,

[1]It is worth mentioning that this interpretation corresponds to the term availability in dependability theory [2].

Manuscript received June 21, 2019; revised November 24, 2019 and March 20, 2020; accepted May 5, 2020. Date of publication May 19, 2020; date of current version August 14, 2020. This research was co-financed by public funding of the state of Saxony/Germany. This work was also supported in parts by the project "Industrial Radio Lab Germany" under contract 16KIS1010K, funded by the Federal Ministry of Education and Research, Germany. The work of E. Jorswieck is partly supported by the German research foundation (DFG) under grant JO 801/24-1. The associate editor coordinating the review of this article and approving it for publication was L. Song. *(Corresponding author: Tom Hößler.)*

Tom Hößler is with the Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, 01069 Dresden, Germany, and also with the Barkhausen Institut, 01187 Dresden, Germany (e-mail: tom.hoessler@tu-dresden.de).

Philipp Schulz and Gerhard P. Fettweis are with the Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, 01069 Dresden, Germany (e-mail: philipp.schulz2@tu-dresden.de; gerhard.fettweis@tu-dresden.de).

Eduard A. Jorswieck is with the Chair for Communications Systems, Technische Universität Braunschweig, 38106 Braunschweig, Germany (e-mail: jorswieck@ifn.ing.tu-bs.de).

Meryem Simsek is with the International Computer Science Institute, Berkeley, CA 94704 USA (e-mail: simsek@icsi.berkeley.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2020.2995150

are discussed with respect to their reliability performance. Analytic comparisons are provided, which build on mathematical proofs. A discussion on which connectivity approach should be selected from each user's perspective is given.
- A novel resource allocation algorithm is introduced, in which a connectivity approach for each user is selected and the maximum number of resources per user is optimized, to satisfy the reliability requirements of all users given a latency budget. Hence, the proposed algorithm performs both, link selection as well as sub-band scheduling over single or multiple wireless links. Relying on matching theory [9], [10], the proposed algorithm is based on stable many-to-one mapping, in which multiple resources are *mapped* to one user, in a multi-user scenario with shared resources.
- In contrast to the pure stable matching, which already exists, our approach additionally combines resource allocation with leveraging the individual optimal connectivity approach for each user and optimizing the maximum number of matched resources, which improves reliability.
- In addition, the proposed resource allocation algorithm is extended by a resource reservation mechanism for weak users (suffering from bad channel conditions), which further enhances the overall reliability performance.
- By means of extensive system-level simulations, we demonstrate the reliability results of the proposed algorithms under different load conditions and different cell densities. It is shown that the achieved performance range can satisfy the strict requirements of URLLC, even in highly loaded scenarios.
- Comparing simulation results in terms of outage probability confirms that both the proposed matching-based algorithm and its resource reservation extension outperform baseline resource allocation approaches, such as Round Robin, Weakest Selects, and random assignment algorithms.

This article is structured as follows: In Section II, existing work related to URLLC in multi-cellular, multi-user systems and stable matching applications in wireless communications are briefly summarized. Section III introduces the system model and presents the problem formulation. In Section IV, the considered connectivity approaches analytically compared, founded on mathematical proofs. Section V recapitulates matching theory basics, followed by the proposed resource allocation algorithm and its extensions. System-level simulation results are discussed in Section VI, before Section VII concludes this article.

II. Related Work

In this section, we summarize research on wireless reliability, focusing on multi-connectivity. Then the concept of matching theory and its application within wireless networks are introduced.

A. Multi-Cellular, Multi-User URLLC

URLLC is considered as one of the key challenges for 5G wireless networks and beyond, receiving major attention from academia and industry. URLLC applications, e.g., wireless factory automation and autonomous driving, combine strict requirements in terms of reliability with latency bounds in the (sub-) millisecond range [11]. Recent advances and diverse challenges of URLLC are reviewed in [12], [13], examining key enablers and their trade-offs with the conclusion that multi-connectivity, among others, is a promising strategy for realizing URLLC.

Multi-connectivity is used as an umbrella term, referring to approaches, where a user equipment (UE) is connected to multiple base stations (BSs). This connection can be on the same or on different frequencies, i.e., intra- and inter-frequency multi-connectivity. Within the context of 5G, different architectural solutions and concepts have been proposed, e.g., [14], [15]. There is a strong trend in research to focus on extremely high data rates by utilizing millimeter wave (mmWave) frequencies provided by multiple BSs (see [16]) and facilitating highly available transmission by combining multiple links, e.g., in [6], [17]. The communication performance of multi-connectivity is quantified in terms of outage probability and throughput in [7]. Concepts of reliability theory have been applied to multi-connectivity scenarios in [18], [19], deriving closed-form expressions for different dependability metrics. In [20] the term "interface diversity" emphasizes the joint utilization of multiple different communication interfaces, which offers additional degrees of diversity and, thus, can help to fulfill the stringent latency-reliability requirements of URLLC.

However, most contributions are restricted to the special case of a single-user scenario. The few contributions on multi-cellular, multi-user evaluations include the following: The system-level performance of multi-user scheduling in 5G is analyzed in [21] without emphasis on URLLC. Reference [22] concludes that fulfilling the URLLC requirements needs novel radio resource management concepts. To the best of our knowledge, there is still a lack of contributions on multi-cellular, multi-user systems which focus on high reliability in combination with low latency. In this paper, we develop a novel analytical framework based on [8], where we demonstrated the feasibility of matching-theory-based multi-connectivity to achieve URLLC requirements for multiple users. Another very recent work in [23] proposes a proactive multi-cell association algorithm and shows how open-loop implementation and multi-cell association enables URLLC.

The following state-of-the-art references on single-connectivity multi-cell multi-user allocation show that the corresponding optimization problems are typically very difficult. Often approximation algorithms are proposed to approach them. In [24], online algorithms for the multi-tier multi-cell user association problem that have provable performance guarantees are proposed based on online combinatorial auctions. A two-sided matching market model is utilized in [25] to

develop an efficient algorithm for user-resource assignments in full-duplex multi-cell networks. The underlying mixed-integer non-linear programming problem is approximated to a geometric problem that is solved by optimality conditions. For multi-cell cooperation in ultra-dense heterogeneous networks, an overview is provided in [26].

B. Matching in Wireless Communications

A situation in which non-divisible goods shall be assigned to entities with different interests can be formulated as a matching problem. One of the most popular matching problems is the stable marriage problem: a set of men and a set of women decide on who to marry based on their preferences over each other, which is a one-to-one matching problem. The notion of stability is important here because it is key for characterizing a robust situation, where no pair of matched partners has an incentive to change the matching. This enables lasting marriages, which are desirable for couples and society at large [27]. Stable matchings have first been studied by *Gale* and *Shapley*, showing that there always exists at least one stable matching, which can be constructed by the so-called deferred acceptance algorithm [27]. Many-to-one stable matchings have numerous applications, e.g. in the labor market and for college admissions [9]. An asymptotic analysis of incentive compatibility and stability in large two-sided matching markets is developed in [28].

In wireless communications, resource allocation problems are central challenges due to the limited resources in time, spectrum, and space [29]. The first comprehensive tutorial on the use of matching theory for resource management in wireless networks is presented in [30]. The authors of [31] discuss the application of matching theory for resource management in wireless networks. [32] provides a comprehensive survey of matching theory, its variants, and their significant properties appropriate for the demands of network engineers and wireless communications. The first application of stable matching in general interference networks is reported in [10]. In heterogeneous networks (HetNets), the assignment of users to their corresponding serving BSs can be modeled as a matching market. In [33], the many-to-one stable matching framework is applied to non-orthogonal spectrum assignment with the goal of maximizing the social welfare of the network. A novel rotation matching algorithm is presented in [34] in order to solve the centralized scheduling and resource allocation problem for a cellular V2X broadcasting system with a focus on access latency. Recently, a many-to-many matching algorithm was proposed aiming to guarantee the reliability requirements of as many users as possible in a multi-cellular, multi-user system in [8], providing a broad overview on wireless multi-connectivity. Resource allocation for URLLC with multiple users based on stable matching is studied in [35], considering a single cell with small-scale fading.

Drawbacks of most of the existing work on stable matching in wireless communications is that either multi-connectivity is not taken into account or only a fixed connectivity approach is utilized. In contrast to previous work available in literature, this article focuses on resource sharing in a multi-cellular, multi-user URLLC system in order to obtain a stable matching with the optimal selection among different connectivity approaches. The literature on matching usually applies fixed quotas, denoting the maximum number of matched partners, as the input to matching procedures. The approach taken in this article is different in this aspect. Instead of presetting fixed quotas, their values are optimized iteratively, which aims for the simultaneous prevention of underprovisioning and starvation of users. In addition, this work proposes a novel extension to matching-based resource allocation, which specifically covers weak users in order to further increase reliability. These contributions complement our prior studies in [8], [10], [35].

III. System Model

This section introduces the deployment scenario, defines different connectivity approaches, and presents the optimization problem.

A. Deployment Scenario and Parameters

We focus on the downlink transmission of a 2-layer HetNet, where layer 1 is modeled as macrocells and layer 2 as small cells. The HetNet consists of the set $\mathcal{M}$ containing $|\mathcal{M}| = M$ hexagonal macrocells overlaid by the set $\mathcal{S}$ of $|\mathcal{S}| = S$ small cells. A BS is either a macrocell eNodeB (MeNB) or a small cell eNodeB (SeNB), i.e. $\mathcal{M} \cap \mathcal{S} = \emptyset$. Within the hexagonal macrocellular area, SeNBs are randomly positioned, so that their coverage areas may overlap. We assume that MeNBs and SeNBs operate in adjacent sub 6-GHz frequencies, whereby SeNBs operate at the same carrier frequency of bandwidth B. This bandwidth B is equally divided into the set $\mathcal{B}$ of $|\mathcal{B}| = N_{\mathrm{B}}$ subbands (SBs). A resource block (RB) is one SB of a single SeNB. This results in a total number of $N = S \cdot N_B$ RBs, each of bandwidth $B_{\mathrm{RB}} = B/N_{\mathrm{B}}$. All RBs are collected in the set $\mathcal{W}$. A set $\mathcal{U}$ of $|\mathcal{U}| = U$ UEs is randomly dropped within the cellular network, whereby a hotspot deployment is considered according to [36].

According to the 3GPP standard, a UE u performs reference signal received power (RSRP) measurements [37]. The MeNB providing the largest RSRP becomes its serving MeNB m_u. Based on a pre-defined timing structure, the UE u sends the RSRP measurements to inform its serving MeNB m_u about its list of potential BSs $\mathcal{C}_u^{\mathrm{pot}}$. This list of potential BSs contains IDs of BSs in a ranked order according to the RSRP values. We assume that the link to MeNB m_u is used for exchanging control information. Especially, UE u's serving MeNB m_u manages connections of UE u to one or more SeNBs based on the RSRP measurements. Allowing connections to several SeNBs extends the concept of dual-connectivity. For the initialization of links to SeNBs, MeNB m_u sends UE u's access requests to potential SeNBs in the set $\mathcal{C}_u^{\mathrm{pot}}$. The set of SeNBs which accept the access request to serve UE u is denoted by $\mathcal{S}_u \subseteq \mathcal{S}$. The set of all cells serving a UE u results as $\mathcal{C}_u = \{m_u\} \cup \mathcal{S}_u$.

The considered traffic model is the URLLC traffic model with periodic packet arrivals defined in [38] under system-level simulation assumptions. In this article, we consider

a fixed number of URLLC traffic UEs with a file size of $F = 200$ bytes and a latency budget of $T_{\text{lat}} = 1\text{ms}$. In this context, ITU and 3GPP discuss URLLC requirements with respect to purely notional packet sizes between 32 bytes and 200 bytes; we select the higher value because it is stricter. The (user plane) latency is defined as the one-way time it takes to successfully deliver an application layer packet/message from the radio protocol layer ingress point to the radio protocol layer egress point of the radio interface in either uplink or downlink in the network for a given service in unloaded conditions, assuming the UE is in the active state [3], [4]. This article does not concentrate on latency optimization. Instead, the focus is on resource sharing for URLLC in a multi-user, multi-cell scenario, taking the required latency into account as a constraint. Thus, the proposed approaches can be easily transferred to different latency values.

The simulated channels take into account path loss, shadowing, and antenna gains, which rely on the sub-6 GHz channel model according to [36]. The link budget of a MeNB m to UE u and SeNB s to UE u in dB is defined as the transmit power minus all losses, which are represented by $L_{u,m}$ and $L_{u,s}$, respectively. The main simulation parameters are summarized in Table I.

B. Connectivity Definitions

In the considered scenario multiple RBs from different SeNBs can be assigned to any UE. In addition, the resulting UE data rate depends on whether the individual RBs of a UE are located in the same SB. The individual connectivity approaches are described in the following.

1) Single-Connectivity (SC): In this case, a UE is connected to only a single SeNB, which is selected based on the measured RSRP. The signal-to-noise ratio (SINR) of UE u, which is connected to *one* SeNB s, is obtained as

$$\gamma_{u,s}^{\text{SC}} = \frac{p_s^{\text{t}} g_{u,s}}{\sum_{s' \in \mathcal{S} \setminus \{s\}} p_{s'}^{\text{t}} g_{u,s'} + \sigma^2}, \tag{1}$$

with p_s^{t} being the transmit power of SeNB s. The propagation gain between UE u and SeNB s is given by $g_{u,s}$, and σ^2 is the noise power.

The propagation gains are defined by $g_{u,s} = 10^{-L_{u,s}/10}$.

UE u's achievable throughput from SeNB s is computed as

$$r_{u,s}^{\text{SC}} = k_{u,s} B_{\text{RB}} \log_2 \left(1 + \gamma_{u,s}^{\text{SC}}\right), \tag{2}$$

with $k_{u,s}$ as the number of RBs assigned to UE u from SeNB s, thus treating interference as Gaussian noise. Each RB has the bandwidth B_{RB}.

2) Multi-Connectivity (MC): In this case, each UE u is assigned to more than one SeNB. We denote the set of assigned SeNB as $\hat{\mathcal{S}}_u \subseteq \mathcal{S}$. We assume that the data/control signals are not transmitted simultaneously from all small cells in $\hat{\mathcal{S}}_u$ since the assigned RBs are located on different SBs. UE u's throughput is defined by

$$r_{u,\hat{\mathcal{S}}_u}^{\text{MC}} = \sum_{s \in \hat{\mathcal{S}}_u} r_{u,s}^{\text{SC}}. \tag{3}$$

This throughput can be achieved by joint decoding, i.e., different but dependent code words are sent via multiple frequency resources and jointly decoded at the receiver [7]. The special case of $|\hat{\mathcal{S}}_u| = 1$ reduces MC to SC.

3) Joint Transmission (JT): In this case, a UE u is connected to more than one SeNB and the assigned RBs are located on the same SBs. We denote the set of assigned SeNB as $\hat{\mathcal{S}}_u^b \subseteq \mathcal{S}$. We assume that the data signals are transmitted simultaneously from all small cells in $\hat{\mathcal{S}}_u$ on the same SB $b \in \mathcal{B}$, which has the bandwidth of one RB B_{RB}. This coordination scheme corresponds to Coordinated Multi-Point JT. We refer to this connectivity approach as JT, assuming that the signal components of the SeNBs fall within the cyclic prefix, resulting in coherent combining of the received signal. The corresponding SINR of UE u in SB b is defined as

$$\gamma_{u,b} = \frac{\sum_{s' \in \hat{\mathcal{S}}_u^b} p_{s'}^{\text{t}} g_{u,s'}}{\sum_{s' \in \mathcal{S} \setminus \hat{\mathcal{S}}_u^b} p_{s'}^{\text{t}} g_{u,s'} + \sigma^2}. \tag{4}$$

UE u's throughput assigned to the set of SeNBs $\hat{\mathcal{S}}_u^b$ over the same SB b is defined by

$$r_{u,b}^{\text{JT}} = B_{\text{RB}} \log_2 \left(1 + \gamma_{u,b}\right). \tag{5}$$

If a UE u is assigned to several SBs, UE u's JT throughput aggregates to

$$r_u^{\text{JT}} = \sum_{b \in \mathcal{B}} r_{u,b}^{\text{JT}}. \tag{6}$$

4) Small Cell Connectivity Generalization: UE u's throughput definition (6) can be utilized to cover all previously introduced connectivity approaches,

$$r_u = \sum_{b \in \mathcal{B}} B_{\text{RB}} \log_2 \left(1 + \gamma_{u,b}\right). \tag{7}$$

Each SB b with no SeNBs assigned to UE u does not contribute to UE u's throughput since $|\hat{\mathcal{S}}_u^b| = 0$ indicates $\gamma_{u,b} = 0$. The set of SBs in which at least one RB is allocated to UE u is referred to as UE u's SBs $\mathcal{B}_u \subseteq \mathcal{B}$. MC is captured if the number of SeNBs assigned to UE u is one for any of UE u's SBs $\mathcal{B}_u$, i.e., $|\hat{\mathcal{S}}_u^b| = 1 \, \forall b \in \mathcal{B}_u$. SC corresponds to the case where a unique SeNB s is assigned to each UE u on all SBs b, i.e., $\hat{\mathcal{S}}_u^b = \{s\} \, \forall b \in \mathcal{B}_u$.

C. Problem Formulation

In the multi-cellular system considered, multiple UEs aim to satisfy their individual service requirement in terms of throughput by optimizing the number of links and the assigned RBs. This optimization is assumed to be performed by the MeNB. Moreover, the resource allocation should ensure stability, i.e., no matched pairs of UEs and RBs have an incentive to swap partners. Due to the limited number of RBs, a minimal resource consumption is targeted. Thus, the optimization problem is formulated as follows:

$$\min_{\mu} \sum_{u \in \mathcal{U}} n_u \tag{8a}$$

subject to:

$$\mu \text{ is a stable matching,} \tag{8b}$$

$$r_u \geq r_u^{\min} \quad \forall u \in \mathcal{U}, \tag{8c}$$

$$p_s^{\text{t}} \leq p_s^{\text{t,max}} \quad \forall s \in \mathcal{S}, \tag{8d}$$

$$\mathcal{S}_u \subseteq \mathcal{C}_u^{\text{pot}} \quad \forall u \in \mathcal{U}. \tag{8e}$$

The objective is to minimize the sum of the numbers of RBs n_u assigned to any UE $u \in \mathcal{U}$, such that the resulting matching is stable (condition (b)). The corresponding definitions are presented in Section V-A. In addition, each UE u's throughput r_u, given by (7), satisfies the minimum throughput requirement $r_u^{\min} := F/T_{\text{lat}}$ by small cell connectivity (condition (c)). Condition (d) implies that the transmit power p_s^{t} of cell c should not exceed the maximum transmit power $p_s^{\text{t,max}}$. Finally, condition (e) guarantees that UE u's serving cells $\mathcal{S}_u$ are selected out the set of potential BSs.

Solving this optimization problem comprises multiple aspects: In the considered multi-user, multi-cellular system, each user throughput depends on the set of serving SeNBs, the number of assigned RBs and the connectivity approach, which specifies the way the RBs are allocated and combined among the frequency SBs. This implies that the optimal connectivity approach for each UE forms the basis for the overall matching outcome of the resource allocation. Thus, we introduce and compare different connectivity techniques in Sec. IV before we present the proposed stable matching procedures in Sec. V, which in turn select the optimal connectivity approach for a given scenario realization built on the analytical findings gained in Sec. IV.

IV. Comparison of Connectivity Approaches

In this section, analytic comparisons between all introduced connectivity approaches are derived. The question to be answered is: "Which connectivity approach should be used from a single user's perspective in order to achieve the highest data rate with a finite number of resources?". Here, it is assumed that the UE's control plane is in the MeNB and the data transmission is performed by the SeNB(s), which are selected according to the RSRP. We assume that all SeNBs transmit on all SBs, corresponding to full frequency reuse.

1) Single-Connectivity vs. Multi-Connectivity: In this section, we compare the performance of SC and MC for a single user.

Theorem 1: For any set $\hat{\mathcal{S}} \subseteq \mathcal{S}$ of SeNBs, which establishes MC to UE u, there exists at least one SeNB $s \in \mathcal{S}$, which achieves at least the same throughput by SC allocating the same number $k = |\hat{\mathcal{S}}|$ of RBs, i.e.,

$$\forall \hat{\mathcal{S}} \subseteq \mathcal{S}\, \exists s \in \mathcal{S} \;\; \textit{such that}\; r_{u,s}^{\text{SC}} \geq r_{u,\hat{\mathcal{S}}_u}^{\text{MC}}. \tag{9}$$

Proof: The maximum throughput for UE u in case of MC in one SB is achieved by assigning the SeNB $s \in \mathcal{S}$ with the highest received power. The same holds true for all SBs. Thus, assigning the SeNB $s \in \mathcal{S}$ with the highest received power for each of the k allocated RBs results in the special case of MC where $\mathcal{S} = \{s\}$ which is equivalent to SC via the SeNB s with $k = k_{u,s}$. Consequently, this SC assignment constitutes an upper bound compared to other MC assignments, which completes the proof. □

This means that a user always has the possibility to waive MC and achieve at least the same performance without the need for more RBs by connecting to only one SeNB. Obviously, this option is preferable in a single user scenario because SC is less cumbersome to realize. However, if multiple UEs are located close to one SeNB, then this selection will lead to congestion.

2) Multi-Connectivity vs. Joint Transmission: In this section, we compare the performance of MC and JT for a single user.

Theorem 2: If one RB of each SeNB in the set $\hat{\mathcal{S}}_u = \hat{\mathcal{S}}_u^b \subseteq \mathcal{S}$, with $|\hat{\mathcal{S}}_u| = |\hat{\mathcal{S}}_u^b| = \hat{S} > 1$, is allocated to a UE u, JT outperforms MC:

$$r_{u,b}^{\text{JT}} > r_{u,\hat{\mathcal{S}}_u}^{\text{MC}}. \tag{10}$$

Proof: See Appendix A. □

Consequently, if RBs of multiple SeNBs are allocated to a user, they should preferably be arranged in the same SB. In this case, a throughput gain due to JT is achieved because interference is turned into useful signal energy. Of course, there is some overhead in joint signal processing, i.e., data fusion and distribution among SeNBs.

3) Single-Connectivity vs. Joint Transmission: In this section, we compare the performance of SC and JT for a single user.

Theorem 3: If $k > 1$ RBs are allocated to UE u, then SC via SeNB $\check{s} \in \mathcal{S}$ outperforms JT via the set $\hat{\mathcal{S}}_u^b$ containing k SeNBs over one SB b, i.e.,

$$r_{u,\check{s}}^{\text{SC}} > r_{u,b}^{\text{JT}}, \quad iff \tag{11}$$

$$p_{\check{s}} > \sqrt[k-1]{\frac{\left(\sum_{\substack{s' \in \mathcal{S} \\ s' \neq \check{s}}} p_{s'} + \sigma^2\right)^k}{\sum_{s' \in \mathcal{S} \setminus \hat{\mathcal{S}}_u^b} p_{s'} + \sigma^2}} - \sum_{\substack{s' \in \mathcal{S} \\ s' \neq \check{s}}} p_{s'} - \sigma^2, \tag{12}$$

where UE u's received powers from all SeNBs are denoted as $p_s = p_s^{\text{t}} g_{u,s}$ with $s \in \mathcal{S}$.

Proof: See Appendix B. □

Unlike the previous connectivity comparisons, it cannot be concluded here that there is always a SC option which offers a higher throughput than JT. The particular result depends on the combination of the actual deployment, path loss, and shadowing with regard to the different SeNBs. However, SC appears to be more powerful in particular situations where the aggregated bandwidth of one SeNB is more valuable than additional weak links. This applies in the case of a dominant SeNB, as illustrated by the following example.

Assume that only one SeNB $\check{s}$ offers a considerable received power from the perspective of a UE u. The received powers from all other SeNBs are, thus, negligible,

$$p_{\check{s}} > 0, \tag{13}$$

$$p_{s'} = 0, \quad \forall s' \in \mathcal{S} \setminus \{\check{s}\}. \tag{14}$$

Hence, eq. (12) reduces to

$$p_{\check{s}} > \sqrt[k-1]{\frac{\sigma^{2k}}{\sum_{s' \in \mathcal{S} \setminus \hat{\mathcal{S}}_u^b} p_{s'} + \sigma^2}} - \sigma^2, \tag{15}$$

If the dominant SeNB is not utilized for JT, $p_{\check{s}} \notin \hat{\mathcal{S}}_u^b$, the JT throughput yields zero, $r_{u,b}^{\text{JT}} = 0$, which is trivial because then SC cannot perform worse than JT. The non-trivial sub-case

$p_{\check{s}} \in \hat{\mathcal{S}}_u^b$, however, results in

$$\sum_{s' \in \mathcal{S} \setminus \hat{\mathcal{S}}_u^b} p_{s'} = 0 \tag{16}$$

due to condition (14). Thus, we obtain

$$r_{u,\check{s}}^{\mathrm{SC}} > r_{u,b}^{\mathrm{JT}}, \quad \text{if} \tag{17}$$

$$p_{\check{s}} > 0 \tag{18}$$

from eq. (15), which shows that SC outperforms JT in cases with a single dominant SeNB.

On the other hand, JT deserves preference in situations where the received powers of the allocated SeNBs are similar to each other, which we demonstrate for the special case of equal received powers, $p_{s'} = p_{\check{s}}\ \forall s' \in \mathcal{S}$.

It follows that

$$\sum_{s' \in \mathcal{S} \setminus \hat{\mathcal{S}}_u^b} p_{s'} = (S-k)p_{\check{s}}, \tag{19}$$

$$\sum_{\substack{s' \in \mathcal{S} \\ s' \neq \check{s}}} p_{s'} = (S-1)p_{\check{s}}, \tag{20}$$

because $|\mathcal{S}| = S$ and $|\hat{\mathcal{S}}_u^b| = k$. Inverting the relation of eq. (11) leads to

$$r_{u,\check{s}}^{\mathrm{SC}} < r_{u,b}^{\mathrm{JT}}, \quad \text{iff} \tag{21}$$

$$\left(p_{\check{s}} + (S-1)p_{\check{s}} + \sigma^2\right)^{(k-1)} < \frac{\left((S-1)p_{\check{s}} + \sigma^2\right)^k}{(S-k)p_{\check{s}} + \sigma^2}. \tag{22}$$

The assumption of received powers which are significantly higher than the noise level,

$$p_{\check{s}} \gg \sigma^2, \tag{23}$$

allows simplifications according to

$$\left(Sp_{\check{s}}\right)^{k-1} < \frac{\left((S-1)p_{\check{s}}\right)^k}{(S-k)p_{\check{s}} + \sigma^2}. \tag{24}$$

In the following, two sub-cases have to be distinguished: If all SeNBs are utilized for JT, $k = S$ implies

$$\left(Sp_{\check{s}}\right)^{S-1} < \frac{\left((S-1)p_{\check{s}}\right)^S}{\sigma^2}. \tag{25}$$

After algebraic manipulations, we obtain that JT outperforms SC in this sub-case of utilizing all SeNBs:

$$r_{u,\check{s}}^{\mathrm{SC}} < r_{u,b}^{\mathrm{JT}}, \quad \text{iff} \tag{26}$$

$$p_{\check{s}} > \frac{\sigma^2 S^{S-1}}{(S-1)^S}, \tag{27}$$

this condition always holds true for the scenario at hand due to the assumption (23). For the other sub-case where JT connections to $k < S$ different SeNBs are established, we rewrite inequality (24) according to

$$0 < (S-1)^k - S^k + kS^{k-1} = \sum_{\ell=0}^{k-2} \binom{k}{\ell} S^\ell (-1)^{k-\ell}, \tag{28}$$

where we waive σ^2 due to assumption (23). In order to demonstrate that this truncated representation of the binomial formula is greater than zero, it is sufficient to focus on consecutive alternating summands and to show that

$$\binom{k}{\ell} S^\ell > \binom{k}{\ell - 1} S^{l-1} \tag{29}$$

because the summand with the highest order is always positive and so each positive summand compensates its negative successor. We apply the definition of the binomial coefficient yielding

$$\frac{k!}{\ell!(k-\ell)!} S^\ell > \frac{k!}{(\ell-1)!(k-l+1)!} S^{\ell-1}, \tag{30}$$

which simplifies to

$$(k - \ell + 1)S = \beta S > \ell. \tag{31}$$

This relation is always satisfied for the considered scenario because $S > k \geq l$ implies $\beta := (k - \ell + 1) \geq 1$. We can conclude that JT via any subset of SeNBs outperforms SC, $r_{u,\check{s}}^{\mathrm{JT}} > r_{u,b}^{\mathrm{SC}}$, if the received powers of the allocated SeNBs are equal. This result is in line with Theorem 1 of our work in [39], where we studied the interplay between the user distribution and performance measures by employing Majorization theory.

V. Proposed Stable Matching Connectivity Algorithm

In this section, we propose an algorithm to achieve a stable matching, which satisfies the formulated optimization problem (8), comprising three components: A variant of the deferred acceptance algorithm is utilized to construct a stable matching. In order to minimize the resource consumption, all the UEs' quotas are optimized, which specify the maximum numbers of allocated RBs per UE. In addition, we propose an extension which guarantees dedicated RBs to the weak UEs if the pure stable matching algorithm is not able to satisfy all UEs.

A. Matching for Resource Allocation

We model the considered scenario as a many-to-one matching game, comprising the two sets of UEs $\mathcal{U}$ and indivisible RBs $\mathcal{W}$ as two teams of players with $\mathcal{U} \cap \mathcal{W} = \emptyset$. RBs can be exclusively assigned to any UE. The UE quota q_u describes how many RBs the UE u can have at most. The problem of assigning the RBs to each UE is a many-to-one matching problem. We assume that all UEs and RBs act independently, i.e., the matching game is a distributed game. Each UE and RB has preferences on the RBs and UEs, respectively. We introduce the notation $a \succ_b \tilde{a}$ meaning that player b prefers player a over player $\tilde{a}$. The corresponding preference lists of UEs l_u^{pref} with $u \in \mathcal{U}$ and RBs l_w^{pref} with $w \in \mathcal{W}$ are obtained based on the SINR values. All players aim for a matching with their most preferred partners. We consider many-to-one matching games focusing on pairwise stability according to the following definitions:

Definition 1: A many-to-one matching μ is a mapping from the set $\mathcal{U} \cup \mathcal{W}$ into the set of all subsets of $\mathcal{U} \cup \mathcal{W}$ such that for each $u \in \mathcal{U}$ and $w \in \mathcal{W}$ the following holds:

1) $\mu(w) \subset \mathcal{U}$ *and* $\mu(u) \subset \mathcal{W}$*;*
2) $|\mu(u)| \leq q_u$*;*
3) $|\mu(w)| \leq q_w = 1$*;*
4) $w \in \mu(u)$ *if and only if* $u \in \mu(w)$*,*

with $\mu(u)$ *(*$\mu(w)$*) being the set of player* u*'s (*w*'s) partners under the matching* μ*.*

Condition 1) describes that players w (u) are matched with players out of the set $\mathcal{U}$ ($\mathcal{W}$). Conditions 2) and 3) guarantee that the number of matched players is at most the same as the players' quota. Condition 4) states that if an RB w is matched to a UE u than this UE u is also matched to the same RB w, which is naturally given in an UE-RB assignment problem.

Definition 2: A blocking pair is the pair of player $u \in \mathcal{U}$ *and player* $w \in \mathcal{W}$*, who prefer each other over some of their partners in the current matching, i.e.,* $u \succ_w \tilde{u}$ *with* $u, \tilde{u} \in \mathcal{U}$ *for some* $\tilde{u} \in \mu(w)$ *and* $w \succ_u \tilde{w}$ *with* $w, \tilde{w} \in \mathcal{W}$ *for some* $\tilde{w} \in \mu(u)$*, respectively.*

Definition 3: The matching is pairwise stable, if there are no blocking pairs.

The definition of stability implies that there is no pair of UE and RB which prefer being matched to each other instead of being matched to their current partner.

B. Deferred Acceptance Algorithm

Every resource allocation problem has at least one stable matching, which can be constructively determined by the so-called deferred acceptance. There is only one stable matching if the preferences of both sets of players are strict and depend on the same metric. The resource-proposing deferred acceptance algorithm yields the stable matching with maximum sum-utility [10]. We utilize this algorithm for the considered resource assignment problem. The corresponding pseudo code is given in Algorithm 1. The matching procedure is assumed to be performed at the MeNB, which receives the required input parameters (preference lists and quotas) from UEs and SeNBs beforehand. The number of iterations of the stable matching algorithm is bounded by the number U of UEs because no RB proposes to a UE twice. After constructing the stable matching, RBs from several SeNBs which are assigned to the same UE are arranged in the same SBs to take advantage of the JT gain, as we derived in Section IV.

C. User Quota Optimization

In conventional matching games, the players' maximum number of partners – their quotas – are assumed to be fixed values which are known before the matching procedure. This is plausible for typical matching problems, e.g., college admissions and applications in the labor market. However, in the context of the considered resource assignment problem, the UE quota q_u, $u \in \mathcal{U}$ is a sensitive hyper-parameter, which should be optimized in order to avoid the following cases: If UE u's quota q_u is too small, the assigned RBs are not sufficient to achieve the required rate, $r_u < r_u^{\min}$. On the other hand, too high quotas may cause over-provisioning of some UEs u, $r_u \gg r_u^{\min}$, if too many resources are assigned. However, this increases the risk of starvation in highly loaded systems, i.e., UEs that have already reached their minimum data rate are assigned further resources instead of unsatisfied UEs. Thus, we propose the following iterative optimization of the individual UE quotas: The initial values are $q_u = 1 \, \forall u \in \mathcal{U}$. After obtaining a stable matching from Algorithm 1, the resulting UE data rates are analyzed. The UE quota is incremented for those UEs whose required minimum rates $r_u^{\min}$ are not achieved by the current stable matching. Then, the stable matching is updated with the improved UE quotas. This procedure is repeated until the stable matching satisfies all UEs' required rates or if the assignment does not change. The latter occurs if all additional potential RBs have rejected UE u's proposal. Then, further increasing UE u's quota q_u has no effect. The amendment of user quota optimization does not weaken the stability because only quotas are changed. Basically, by choosing decent quotas, the stable matching outcome is steered towards the desired assignment.

Algorithm 1 Many-to-One Stable Matching

Input: UE quotas q_u, preference lists of all UEs l_u^{pref} with $u \in \mathcal{U}$ and all RBs l_w^{pref} with $w \in \mathcal{W}$

Proposing and Matching:

Step $t = 0$*:*

Initialize the ordered set of UE u's temporarily accepted RBs $\mathcal{A}^t(u) = \emptyset$ for $u \in \mathcal{U}$.

Step t*:*

Proposals:

Every RB not yet assigned $w \in \mathcal{W} \setminus \bigcup_{u \in \mathcal{U}} \mathcal{A}^{t-1}(u)$ sends a proposal to its most preferred UE $u \in \mathcal{U}$ (via its MeNB). This index is cleared from the preference list l_w^{pref} of RB w.

Decisions:

Denote RBs which proposed to UE u in step t as $\mathcal{P}^t(u)$. UE u keeps the q_u best ranked RBs from $\mathcal{A}^{t-1} \cup \mathcal{P}^t(u)$ with subject to its preference list l_u^{pref} and updates $\mathcal{A}^t(u)$ accordingly.

Output: Stable matching μ

D. Resource Reservation

Especially in highly loaded systems, it is possible that some UEs remain unsatisfied by the stable matching. Weak UEs are detected as those whose throughput resulting from the stable matching does not meet the required threshold. In these cases, we propose to exclude those weak UEs from the stable matching procedure. Instead, the weak UEs are allowed to allocate their most preferred RBs, choosing the optimal connectivity approach according to the findings in Section IV. Subsequently, the stable matching algorithms is performed with respect to the remaining UEs and RBs. The combination of the stable matching and the resource reservation leads to the (temporary) allocation. Due to the reduced number of RBs which are available for stable matching, further UEs may become unsatisfied. In this case, they are considered for resource reservation, as well. This procedure is repeated until all UEs' required rates are satisfied or more than a whole frequency band is requested for resource reservation. In the latter case the temporary allocation result is applied which yields the minimal number of unsatisfied UEs. The UEs who

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
Cellular layout	Hexagonal grid, 3 sectors per cell	Max. MeNB (SeNB) transmit power	46 dBm (30 dBm)
Number of MeNBs	1	MeNB path loss	$128.1 + 37.6 \log_{10}(d/\text{km})$ dB
Number of SeNBs per sector	$\{1; 3; 5; 10; 15; 20\}$	SeNB path loss	$140.7 + 36.7 \log_{10}(d/\text{km})$ dB
Carrier frequency	2 GHz	Thermal noise density	-174 dBm/Hz
Bandwidth	20 MHz	Shadowing std.	10 dB
Number of SBs	100	MeNB (SeNB) antenna gain	14 dBi (5 dBi)
Subframe duration	1 ms	Transmission mode	2×2 MIMO
Traffic model	Periodic URLLC traffic	Min. dist. SeNBs - SeNB (MeNB)	40 m (75 m)
Required min. UE rate	1.6 Mbps	Min. dist. SeNB - UE	10 m
Sum of simulated UEs	3 000 000	Number of hotspot UEs	$\lceil 2/3U \rceil$
RSRP threshold	-114 dBm	Hotspot radius	40 m

are satisfied by resource reservation meet their requirements and are excluded from the stable matching procedure. They therefore do not compromise the stability of the matching between the remaining UEs and RBs.

E. Complexity

In this section, we analyze the complexity of the presented algorithms, utilizing the big O notation, which provides an asymptotic upper bound on the corresponding growth rates.

1) Many-to-One Stable Matching: The preference lists are established based on $U \cdot S$ RSRP measurements between all UEs and SeNBs. During the initialization step ($t = 0$), each UE determines the temporarily accepted RBs, yielding $\mathcal{O}(U)$. The complexity of the subsequent proposal and decision phase ($t > 0$) can be expressed by $\mathcal{O}(N \cdot \max\{q_u\})$. The matching algorithm is guaranteed to terminate in a matching after $t \leq U$ steps because no RB proposes to a UE twice. This results in the complexity $\mathcal{O}(U \cdot \max\{U, N \cdot \max\{q_u\}\}) = \mathcal{O}(UN \cdot \max\{q_u\})$ under the reasonable assumption of less UEs than RBs, $U < N$.

2) User Quota Optimization: Every UE has the initial user quota of $q_u = 1 \, \forall u \in \mathcal{U}$. There can be at most $N - U$ quota increments in total, since there are only $N - U$ resources available after an initialization of one RB per UE. Thus, optimizing the user quota is carried out in complexity $\mathcal{O}(N - U) = \mathcal{O}(N)$ with $U < N$.

3) Resource Reservation: In the worst case, each iteration of resource reservation leads to a single new unsatisfied UE due to the reduced number of RBs which are left for stable matching. This corresponds to complexity $\mathcal{O}(U)$.

Finally, the overall complexity including user quota optimization and resource reservation results by concatenating the three algorithms according to $\mathcal{O}(UN \cdot \max\{q_u\}) \cdot [\mathcal{O}(N) + \mathcal{O}(U)] = \mathcal{O}(N^2 U \cdot \max\{q_u\})$.

VI. SIMULATION RESULTS

In this section, we introduce the parameters of the system-level simulation scenario and present performance evaluation results for single-user and multi-user settings.

A. Simulation Scenario

In this section, the proposed solutions are validated in a system-level simulator based on the assumptions and parameters defined in [36]. We consider a HetNet with a macrocell consisting of $S/3 \in \{1, 3, 5, 10, 15, 20\}$ small cells per macro sector, uniformly and randomly distributed within the macrocellular environment. Two thirds of U UEs are randomly and uniformly dropped within a 40 m radius of each small cell s. The remaining UEs are uniformly distributed within the macrocellular area. Periodic URLLC traffic is considered for each UE. Each UE has a file size of $F = 200$ bytes to be downloaded within a latency budget of 1 ms i.e., the minimum required data rate is $r^{\min} = r_u^{\min} = 1.6\text{Mbps} \ \forall u \in \mathcal{U}$. Our system-level simulation results are determined over 3 000 000 random realizations. Further details about the system-level simulation parameters are provided in Table I. The results are not bounded by a maximum modulation and coding scheme because the bounding affects extremely high throughput values, which are out of scope for URLLC.

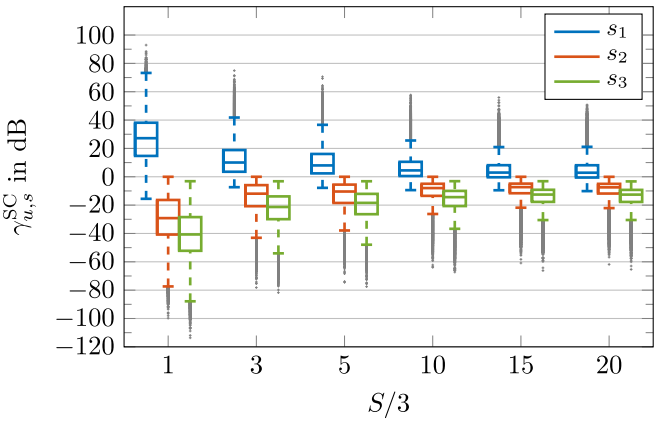


Fig. 1. SINR to the three strongest SeNBs, denoted as s_1, s_2, s_3, for different numbers of SeNBs per sector.

B. Single-User Performance

At first, we focus on a single-user scenario in order to investigate the impact of the small-cell deployment on the user's SINR and data rate. In Fig. 1, we present box plots of the three strongest SeNBs with respect to the SINR from each UE's perspective comparing different SeNB densities. The bottom and top of a box are the first and third quartiles, respectively. The band inside the box is the median, outliers

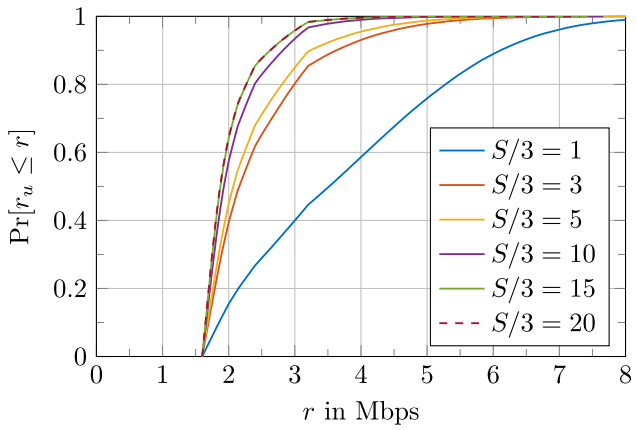


Fig. 2. CDFs of single user rates r_u for different numbers of SeNBs per sector.

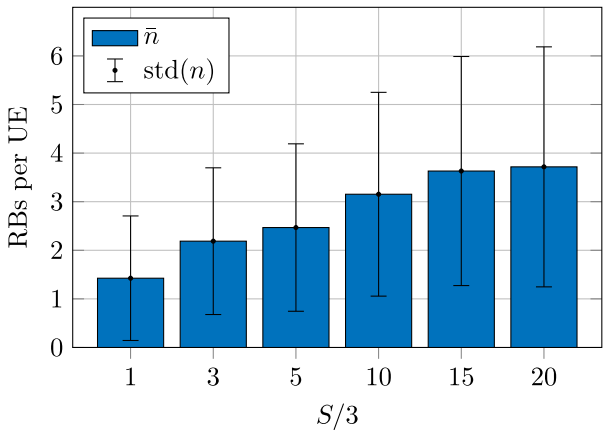


Fig. 3. RBs per UE for different numbers of SeNBs per sector.

are depicted as gray markers. As expected, the SINR to the strongest SeNB decreases for a higher number of SeNBs due to increasing interference. On the other hand, the SINR regarding the second and third strongest SeNBs increase at a higher density of SeNBs. For high numbers of SeNBs, the SINR variances are also reduced for each SeNB rank as well as the aggregated data. The results of the two cases with highest SeNB densities, $S/3 = \{15; 20\}$, are almost equivalent to each other. The SINR values of the second and third strongest SeNBs are not higher than 0dB, which directly results from the SINR definition.

Based on these SINR values, the proposed stable matching algorithm is performed. The resulting user rates for the considered SeNB densities are depicted as cumulative distribution functions (CDFs) in Fig. 2. Since a single user scenario is considered here, each UE is matched to its most preferred SeNB, which is the strongest SeNB in terms of SINR. It can be observed that all rates are higher than 1.6Mbps, satisfying the required minimum rate. So the maximum reliability is achieved in the single user scenario. The quota optimization procedure ensures that no fewer RBs are allocated to a UE than required. At the same time, this aims to prevent overprovisioning of a UE. For any percentile, the rate increases for smaller SeNB densities. This is due to the fact that the strongest SeNB provides a higher SINR in less dense deployments, implying higher throughput per RB because the bandwidth of an RB is fixed. The differences between the two cases of highest SeNB density, $S/3 = \{15; 20\}$, are not visible, complying with the observation of their SINR values.

The plot in Fig 3 shows the average number and the standard deviation of RBs a UE is allocated to, in order to satisfy the required rate. Both values increase for higher SeNB densities because of more severe interference. From a single user's perspective, the least dense SeNB deployment scenario seems to be preferable because the required rate can be surpassed the most, utilizing the least resources. However, this is only true if any UE can be served by its most preferred SeNB and if enough RBs are available, such that no competition for resources is fought. Of course, this is fundamentally different in a multi-user scenario, which is more relevant for practical systems.

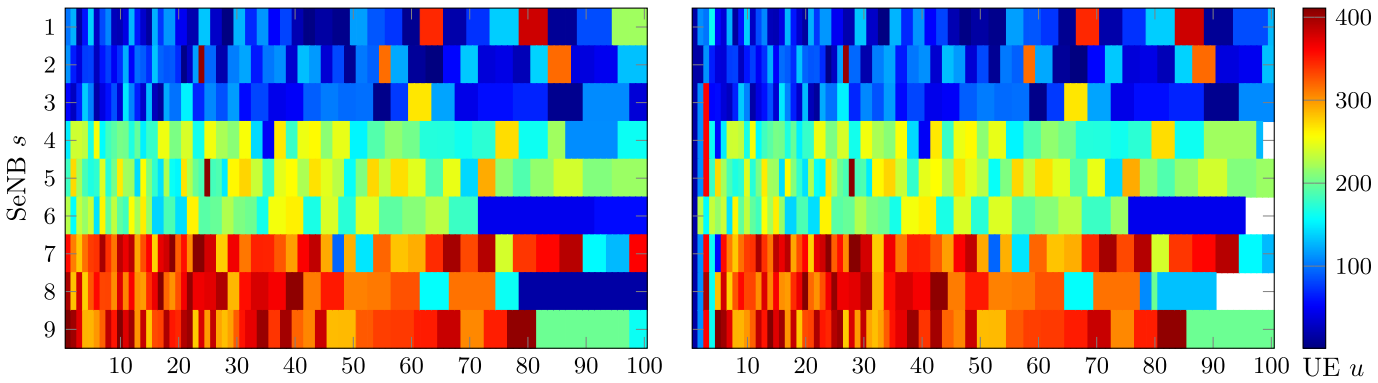
TABLE II
OPERATING POINTS: NUMBERS OF UES PER SECTOR, $U/3$

$S/3$	Medium load (75%)	Full load (100%)	Overload (125%)
1	52	70	87
3	102	137	171
5	151	202	252
10	237	317	396
15	309	413	516
20	402	537	671

C. Multi-User Performance

We evaluate system-level simulations for several multi-user scenarios to demonstrate the performance of the proposed stable matching resource allocation if UEs compete for a limited number of RBs. In order to compare different loads in terms of numbers of UEs, the average numbers of RBs per UE from the single-user scenario are extrapolated. We define full load as the number of UEs expected to require all RBs. In addition, medium load (75%) and overload (125%) are investigated. The corresponding operating points for all considered numbers of SeNB per sector are summarized in Table II. $3\,000\,000$ simulations have been conducted for each operating point and resource allocation approach.

In order to demonstrate the performance of the proposed stable matching algorithm and the extension of resource reservation, Fig. 4 visualizes exemplary resource allocations for one full load realization with $S = 9$ SeNBs. Each row contains the RBs of one SeNB, the columns correspond to SBs. The UE index, a RB is allocated to, is reflected by the color. Without loss of generality, the first/second/third 137 UEs are randomly dropped in sector 1/2/3, which is reflected by shades of blue, green, and red color. Analogously, the SeNBs are distributed to the three sectors. Fig. 4a shows the resource allocation resulting from the proposed stable matching algorithm. It can be seen that all RBs are allocated. Due to the quota optimization, the number of allocated RBs differs among the UEs. In general, the number of RBs allocated to a certain UE is low if the UE is matched to an SeNB which offers a high SINR. Allocations of this type dominate the left part of the plot, because the corresponding proposals of RBs are accepted



(a) Stable matching. (b) Stable matching with resource reservation.

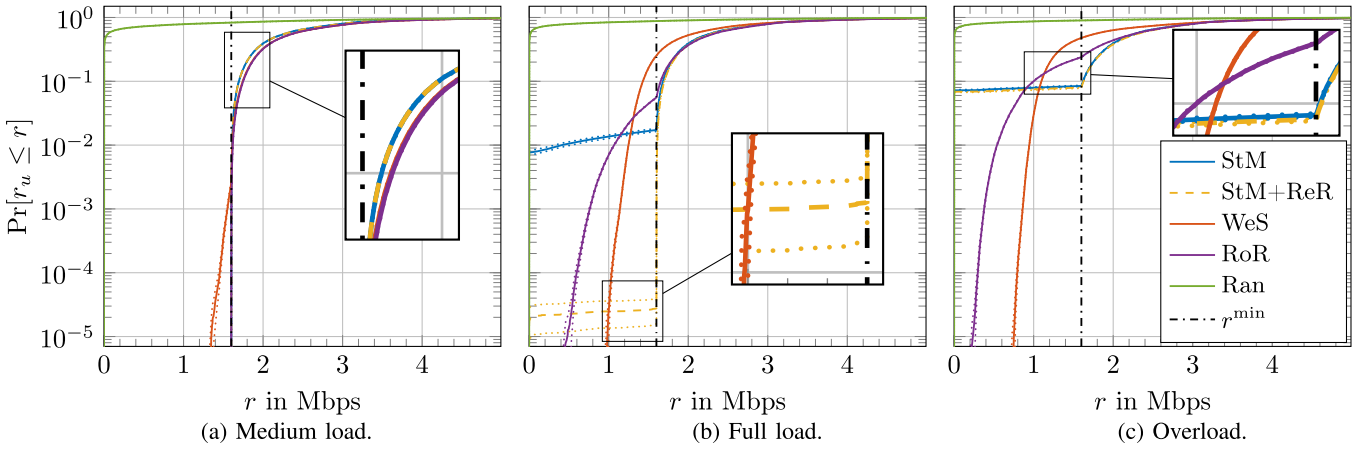
Fig. 4. Exemplary resource allocation for three SeNBs per sector at full load achieved by (a) stable matching and (b) stable matching with resource reservation. The color bar in (b) also applies to (a).

by the UEs as part of the stable matching algorithm. The RBs of a SeNB share the same preference lists because preferences of RBs among the UEs only depend on the SINR value to their SeNB. Due to the competition for limited resources, not every UE can be matched to RBs of the most preferred SeNB. This results in a higher number of requested RBs, equivalent to a higher quota, which is illustrated by wider boxes of the same color. In this particular example, the four UEs assigned to last RBs of SeNBs 3, 6, 7, and 8 are not able to satisfy the required rate, despite of increased quotas. Thus, the proposed resource reservation with respect to those weak UEs is executed. The allocation matrix, obtained after two iterations, is presented in Fig. 4b. The guaranteed resources for weak UEs, which are visible in parts of the first columns, ensure that all UEs satisfy the required rate. Thus, some RBs remain unused, depicted as white areas. The resource reservation procedure selects the best connectivity approach, taking advantage of the analytical findings in Sec. IV. In case of UEs assigned to several SeNBs, the corresponding RBs are aligned among the SBs to exploit the JT gain, discussed in Sec. IV-.2. Thus, the resource reservation does not utilize complete SBs evenly, e.g., in Fig. 4b SeNBs 4, 5, and 6 reserve 5, 3, and 1 RBs, respectively. This effect creates some offsets in the RBs, assigned by stable matching, which are not in the same trend, comparing both resource allocations in Fig. 4. Besides that, the allocation is very similar to that obtained by pure stable matching because all UEs except for the weak UEs perform the stable matching algorithm. However, the following tradeoff can be observed: Those RBs which are reserved to the weak UEs are no longer available for the stable matching procedure. Hence, some of the previously satisfied UEs are matched to different SeNBs, which they do not prefer over their initial allocation. They in turn need other resources to meet the minimum required rate $r^{\min}$. For instance, the (green) UE 219 is satisfied by the stable matching with the assignment to SeNBs 1 and 4 on SBs 95 – 100 and 84 – 86, respectively. However, after resource reservation, other UEs occupy SBs 95 – 100 of SeNB 1, who are preferred by the SeNB. Consequently, UE 219 is instead assigned to other RBs from SeNB 4, i.e., on SBs 89 – 97.

In order to extend the performance evaluation of the proposed algorithms, we present the CDFs of user rates for $S = 9$ SeNBs for different loads in Fig. 5. Our proposed approaches stable matching (StM) and stable matching with resource reservation (StM + ReR) are benchmarked to the following allocation algorithms:

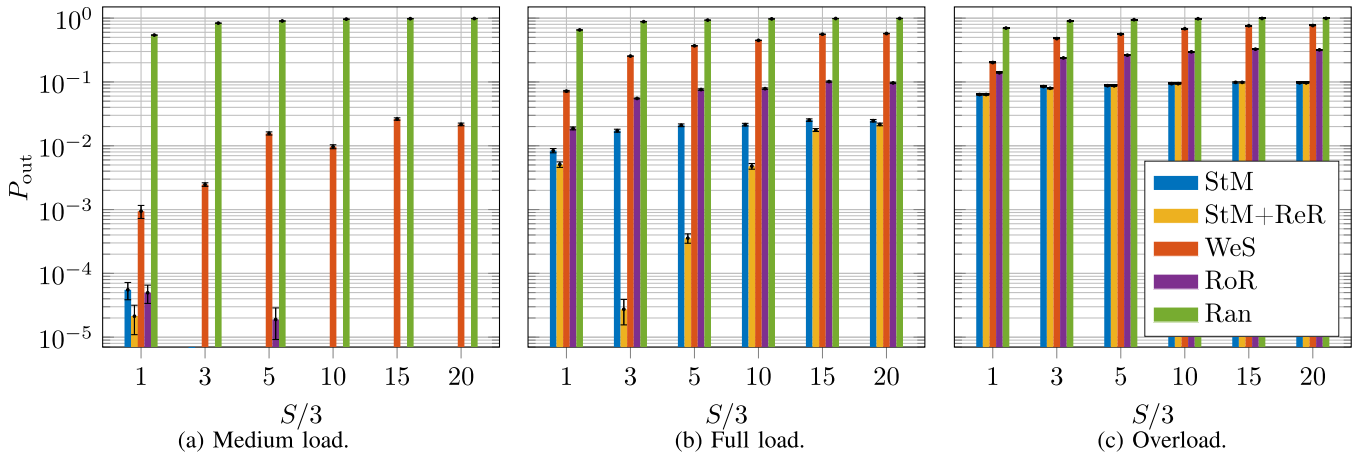
- adaptive resource allocation presented in [40], which we refer to as "Weakest Selects" (WeS),
- Round Robin allocation (RoR),
- random allocation (Ran).

The required rate $r^{\min} = 1.6\text{Mbps}$ is included to represent the targeted threshold. The included confidence bounds of 99.99% (dotted lines) are determined by Greenwood's formula [41]. The fact that the confidence bounds are very tight indicates a low uncertainty of the obtained simulation results, even in the range of the plotted low percentiles. The probability of not achieving the required rate is denoted as outage probability P_{out}. In the CDF, it corresponds to the percentile of $r = r^{\min}$. The optimal CDF would be a step function from zero to one at the required rate, corresponding to the case that the required rate is met by every UE. In Fig. 5b, it can be seen that the rates obtained by the compared algorithms differ significantly. StM outperforms Ran, WeS, and RoR in terms of outage probability. Performing StM + ReR further improves the outage probability by almost three orders of magnitude, resulting in the range of $4 \cdot 10^{-5}$, which is relevant to URLLC use cases. All algorithms except for Ran take the targeted rate into account. However, notches in the CDF curves only appear for RoR and the proposed matching algorithms, approaching the ideal result of a step function. The extent of the notches characterizes the sensitivity against the minimum required rate. Any rate below the minimal requirement $r^{\min}$ corresponds to an outage, which is reasonable for periodic URLLC traffic. Hence, flat slopes of the CDFs are desirable for $r < r^{\min}$, because it is preferable from the perspective of the multi-user system that a UE unable to reach the target throughput releases its resources to other users. The different slopes below the threshold $r^{\min}$ reflect the various ranges of rates for unsatisfied users. The crossing points of multiple curves, however, are not relevant for further performance evaluation in the context of



(a) Medium load. (b) Full load. (c) Overload.

Fig. 5. CDFs of user rate r_u with three SeNBs per sector at different loads. The dotted lines indicate the 99.99% confidence bounds. The legend of (c) also applies to (a) and (b).



(a) Medium load. (b) Full load. (c) Overload.

Fig. 6. Outage probability for three SeNBs per sector at different loads. The error bars represent the 99.99% confidence interval. The legend of (c) also applies to (a) and (b).

URLLC because they are located below the required minimum rate. Fig. 5a shows that for medium load, all resource allocation algorithms improve their outage probability compared to full load, except for Ran. It can be observed that both proposed algorithms, StM and StM + ReR, as well as RoR achieve outage probabilities below 10^{-5}. However, WeS is not able to satisfy all UEs although 25% of RBs are not requested on average in the medium load case. As expected, in overloaded scenarios, all resource allocation algorithms perform worst, as visualized in Fig. 5c. The proposed stable-matching-based approaches still outperform all reference algorithms. However, resource reservation can only slightly improve the UE rates due to the lack of resources. In all CDFs, it is visible that the rate of some users greatly exceeds the minimum rate constraint. However, this does not imply that the number of RBs can be further minimized, because the considered RBs are not continuous but discrete resources. Thus, in case of users who do not meet the rate constraint tightly, a single additional allocated RB may lead to exceeding the threshold by far. On the other hand, it is possible, that users with excellent channel conditions greatly exceed the minimum required rate with a single RB.

After concentrating on one selected deployment scenario, we compare the outage probabilities of all considered resource allocation algorithms for different small cell densities and several user loads as depicted in Fig. 6. For full load conditions, Fig. 6b shows that the proposed algorithms, StM and StM + ReR, outperform the reference approaches for each considered SeNB density, generalizing the previous observations. The gain through resources reservation differs: Obviously, the overall minimal outage probability is achieved by StM + ReR for three SeNBs per sector. Besides that, the outage probabilities seem to generally deteriorate with respect to higher numbers of SeNBs. RoR is the reference algorithms, which performs best, followed by WeS. Regarding medium load, it is clearly visible in Fig. 6a that except for the deployment with one SeNB per sector, the proposed resource allocation algorithms achieve outage probabilities lower than 10^{-5}, RoR performs nearly as well. The reason for higher outage probabilities of the proposed stable-matching-based algorithms in the least dense scenario is the fact that the SINR provided by others than the strongest SeNB are weaker compared to the other deployment scenarios. Thus, a UE not matched to RBs of its most preferred SeNB would request

a high number of RBs from other SeNB, which cannot be provided in all cases – even at medium load. Fig. 6c illustrates the resulting outage probabilities at overload. Again, the proposed algorithms outperform the reference resource allocation strategies. However, the limits of resource reservations are obvious since it hardly improves the outage probability of the pure StM. Similar to the full load case, the reference approaches can be sorted according to their ascending outage probability as follows: RoR, WeS, Ran.

VII. Conclusion

In this article, we have studied multi-cellular, multi-user heterogeneous cellular systems in which all users have the same stringent URLLC requirements. We proposed a novel matching-based resource allocation algorithm, supporting multi-connectivity which aims for satisfying the throughput requirement of all users by optimizing quotas, i.e., the maximum numbers of RBs allowed to be assigned per UE. The presented resource reservation extension for weak users is able to significantly enhance reliability of the whole system. Our system-level simulations for different load and density conditions demonstrate that the proposed resource allocation algorithms outperform state-of-the-art approaches. Analytic comparisons between different small-cell connectivity approaches reveal that multi-connectivity may not be always optimal in the considered scenarios. Thus, we provide novel insights on how to carefully optimize the resource allocation for wireless URLLC, coping with challenges like interference and competition for limited resources. The proposed algorithms result in enhanced reliability satisfying the targeted performance of wireless URLLC. Consequently, this article will help to foster research on URLLC, whose importance will remain significant even beyond 5G.

Acknowledgment

The authors would like to thank the Center for Information Services and High Performance Computing (ZIH) at TU Dresden for generous allocations of computer time.

Appendix A
Proof of Theorem 2

Applying the definitions (1) to (6) to Theorem 2 and applying algebraic manipulations yield the equivalent expressions

$$B_{\text{RB}} \log_2 \left(1+\gamma_{u,b}\right) > \sum_{s\in\hat{\mathcal{S}}_u} B_{\text{RB}} \log_2 \left(1+\gamma^{\text{SC}}_{u,s}\right) \quad (32)$$

$$1+\frac{\sum_{s\in\hat{\mathcal{S}}_u} p^{\text{t}}_s g_{u,s}}{\sum_{s'\in\mathcal{S}\setminus\hat{\mathcal{S}}_u} p^{\text{t}}_{s'} g_{u,s'}+\sigma^2} > \prod_{s\in\hat{\mathcal{S}}_u}\left(1+\frac{p^{\text{t}}_s g_{u,s}}{\sum_{\substack{s'\in\mathcal{S}\\ s'\neq s}} p^{\text{t}}_{s'} g_{u,s'}+\sigma^2}\right) \quad (33)$$

Without loss of generality the following notation is introduced to improve readability. UE u's received powers are denoted by $p_j = p^{\text{t}}_{s_j} g_{u,s_j}$ with $s_j \in \hat{\mathcal{S}}_u$ and $j = 1, 2, \ldots, \hat{S}$. UE u's interference powers and noise power is collected as

$$p_{\text{in}} = \sum_{s'\in\mathcal{S}\setminus\hat{\mathcal{S}}_u} p^{\text{t}}_{s'} g_{u,s'} + \sigma^2. \quad (34)$$

Thus, Theorem 2 is equivalent to the following Proposition.

Proposition 1: Let $\hat{S} \in \mathbb{N}\setminus\{0,1\}$ and $p_{\text{in}} > 0$. Then for all $p_1, p_2, \ldots, p_{\hat{S}} > 0$ the following inequality holds

$$\frac{p_{\text{in}} + \sum_{i=1}^{\hat{S}} p_i}{p_{\text{in}}} > \prod_{j=1}^{\hat{S}} \frac{\sum_{i=1}^{\hat{S}} p_i + p_{\text{in}}}{\sum_{i=1}^{\hat{S}} p_i - p_j + p_{\text{in}}}. \quad (35)$$

Proof: [Proof by induction] For $\hat{S} = 2$ Eq. (35) holds, since

$$0 < p_1\ p_2 \quad (36)$$

$$p_1\ p_{\text{in}} + p_2\ p_{\text{in}} + p_{\text{in}}^2 < p_1 p_2 + p_1\ p_{\text{in}} + p_2\ p_{\text{in}} + p_{\text{in}}^2 \quad (37)$$

$$(p_1 + p_2 + p_{\text{in}})p_{\text{in}} < (p_1 + p_{\text{in}})(p_2 + p_{\text{in}}) \quad (38)$$

$$\frac{1}{(p_1 + p_2 + p_{\text{in}})p_{\text{in}}} > \frac{1}{p_1 + p_{\text{in}}} \frac{1}{p_2 + p_{\text{in}}} \quad (39)$$

$$\frac{p_1 + p_2 + p_{\text{in}}}{p_{\text{in}}} > \frac{p_1 + p_2 + p_{\text{in}}}{p_1 + p_{\text{in}}} \frac{p_1 + p_2 + p_{\text{in}}}{p_2 + p_{\text{in}}}. \quad (40)$$

Assuming that Eq. (35) holds for $\hat{S} = N$ implies that it also holds for $\hat{S} = N + 1$ as shown in the following. The right hand side of Eq. (35) yields

$$Z := \prod_{j=1}^{N+1} \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_j + p_{\text{in}}} \quad (41)$$

$$= \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_{N+1} + p_{\text{in}}} \prod_{j=1}^{N} \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_j + p_{\text{in}}} \quad (42)$$

$$= \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_{N+1} + p_{\text{in}}} \prod_{j=1}^{N} \frac{\sum_{i=1}^{N} p_i + p_{N+1} + p_{\text{in}}}{\sum_{i=1}^{N} p_i - p_j + p_{N+1} + p_{\text{in}}}. \quad (43)$$

With $\hat{p} = p_{N+1} + p_{\text{in}} > 0$, the inductive assumption is used for the second factor,

$$Z < \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_{N+1} + p_{\text{in}}} \frac{\sum_{i=1}^{N} p_i + \hat{p}}{\hat{p}} \quad (44)$$

$$= \frac{\sum_{i=1}^{N+1} p_i + p_{\text{in}}}{\sum_{i=1}^{N+1} p_i - p_{N+1} + p_{\text{in}}} \frac{\sum_{i=1}^{N} p_i + p_{N+1} + p_{\text{in}}}{p_{N+1} + p_{\text{in}}} \quad (45)$$

$$= \frac{\sum_{i=1}^{N} p_i + p_{N+1} + p_{\text{in}}}{\sum_{i=1}^{N} p_i + p_{\text{in}}} \frac{\sum_{i=1}^{N} p_i + p_{N+1} + p_{\text{in}}}{p_{N+1} + p_{\text{in}}} \quad (46)$$

$$= \frac{\hat{p_1} + \hat{p_2} + p_{\text{in}}}{\hat{p_1} + p_{\text{in}}} \frac{\hat{p_1} + \hat{p_2} + p_{\text{in}}}{\hat{p_2} + p_{\text{in}}}. \quad (47)$$

With $\hat{p_1} = \sum_{i=1}^{N} p_i > 0$ and $\hat{p_2} = p_{N+1} > 0$, the exact stucture of Eq. (35) for $\hat{S} = 2$ (c.f. Eq. (40)) is obtained,

$$Z < \frac{p_{\text{in}} + \hat{p_1} + \hat{p_2}}{p_{\text{in}}} = \frac{p_{\text{in}} + \sum_{i=1}^{N+1} p_i}{p_{\text{in}}}. \quad (48)$$

This completes the proof by the principle of induction. □

Appendix B
Proof of Theorem 3

Proof: Applying the definitions (1), (2), (4), (6) to Eq. (11) of Theorem 3 and applying algebraic manipulations yield

$$k B_{\text{RB}} \log_2 \left(1+\gamma^{\text{SC}}_{u,\check{s}}\right) > B_{\text{RB}} \log_2 \left(1+\gamma_{u,b}\right) \quad (49)$$

$$\left(1+\frac{p_{\check{s}}}{\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2}\right)^k > 1+\frac{\sum_{s'\in\hat{\mathcal{S}}_{u,b}} p_{s'}}{\sum_{s'\in\mathcal{S}\setminus\hat{\mathcal{S}}_{u,b}} p_{s'}+\sigma^2} \tag{50}$$

$$\left(\frac{p_{\check{s}}+\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2}{\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2}\right)^k > \frac{p_{\check{s}}+\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2}{\sum_{s'\in\mathcal{S}\setminus\hat{\mathcal{S}}_{u,b}} p_{s'}+\sigma^2} \tag{51}$$

$$\left(p_{\check{s}}+\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2\right)^{k-1} > \frac{\left(\sum_{\substack{s'\in S\\ s'\neq\check{s}}} p_{s'}+\sigma^2\right)^k}{\sum_{s'\in\mathcal{S}\setminus\hat{\mathcal{S}}_{u,b}} p_{s'}+\sigma^2}. \tag{52}$$

Solving for $p_{\check{s}}$ results in Eq. (12), which completes the proof. □

REFERENCES

[1] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile Internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 460–473, Mar. 2016.

[2] T. Hosler, L. Scheuvens, N. Franchi, M. Simsek, and G. P. Fettweis, "Applying reliability theory for future wireless communication networks," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–7.

[3] *Minimum Requirements Related to Technical Performance For IMT-2020 Radio Interface(s)*, document M.2410-0, ITU-R, Nov. 2017.

[4] *Study on Scenarios and Requirements For Next Generation Access Technologies, (V14.3.0)*, document 38.913, 3GPP, Oct. 2017.

[5] D. Ohmann, M. Simsek, and G. P. Fettweis, "Achieving high availability in wireless networks by an optimal number of Rayleigh-fading links," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, USA, Dec. 2014, pp. 1402–1407.

[6] D. Ohmann, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Achieving high availability in wireless networks by inter-frequency multi-connectivity," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–7.

[7] A. Wolf, P. Schulz, M. Dörpinghaus, J. C. S. S. Filho, and G. Fettweis, "How reliable and capable is multi-connectivity?" *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1506–1520, Feb. 2019.

[8] M. Simsek, T. Hobler, E. Jorswieck, H. Klessig, and G. Fettweis, "Multiconnectivity in multicellular, multiuser systems: A Matching-based approach," *Proc. IEEE*, vol. 107, no. 2, pp. 394–413, Feb. 2019.

[9] A. E. Roth and M. A. O. Sotomayor, *Two-Sided Matching—A Study in Game-Theoretic Modeling and Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[10] E. A. Jorswieck, "Stable matchings for resource allocation in wireless networks," in *Proc. 17th Int. Conf. Digit. Signal Process. (DSP)*, Corfu, Greece, Jul. 2011, pp. 1–8.

[11] P. Schulz *et al.*, "Network architectures for demanding 5G performance requirements: Tailored toward specific needs of efficiency and flexibility," *IEEE Veh. Technol. Mag.*, vol. 14, no. 2, pp. 33–43, Jun. 2019.

[12] N. A. Mohammed, A. M. Mansoor, and R. B. Ahmad, "Mission-critical machine-type communication: An overview and perspectives towards 5G," *IEEE Access*, vol. 7, pp. 127198–127216, 2019.

[13] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[14] D. S. Michalopoulos, I. Viering, and L. Du, "User-plane multi-connectivity aspects in 5G," in *Proc. 23rd Int. Conf. Telecommun. (ICT)*, Thessaloniki, Greece, May 2016, pp. 1–5.

[15] A. Ravanshid *et al.*, "Multi-connectivity functional architectures in 5G," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 187–192.

[16] M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Multi-connectivity in 5G mmWave cellular networks," in *Proc. Medit. Ad Hoc Netw. Workshop (Med-Hoc-Net)*, Vilanova i la Geltrú, Spain, Jun. 2016, pp. 1–7.

[17] D. Ohmann, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Diversity trade-offs and joint coding schemes for highly reliable wireless transmissions," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Montreal, QC, Canada, Sep. 2016, pp. 1–6.

[18] T. Hoesler, M. Simsek, and G. P. Fettweis, "Mission reliability for URLLC in wireless networks," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2350–2353, Nov. 2018.

[19] T. Hossler, M. Simsek, and G. P. Fettweis, "Joint analysis of channel availability and time-based reliability metrics for wireless URLLC," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.

[20] P. Popovski *et al.*, "Wireless access in ultra-reliable low-latency communication (URLLC)," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5783–5801, Aug. 2019.

[21] K. Pedersen, G. Pocovi, J. Steiner, and A. Maeder, "Agile 5G scheduler for improved E2E performance and flexibility for different network implementations," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 210–217, Mar. 2018.

[22] G. Pocovi, H. Shariatmadari, G. Berardinelli, K. Pedersen, J. Steiner, and Z. Li, "Achieving ultra-reliable low-latency communications: Challenges and envisioned system enhancements," *IEEE Netw.*, vol. 32, no. 2, pp. 8–15, Mar. 2018.

[23] C.-H. Liu, D.-C. Liang, K.-C. Chen, and R.-H. Gau, "Ultra-reliable and low-latency communications using proactive multi-cell association," 2019, *arXiv:1903.09739*. [Online]. Available: http://arxiv.org/abs/1903.09739

[24] W. C. Ao and K. Psounis, "Approximation algorithms for online user association in multi-tier multi-cell mobile networks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 4, pp. 2361–2374, Aug. 2017.

[25] S. Sekander, H. Tabassum, and E. Hossain, "Decoupled uplink-downlink user association in multi-tier full-duplex cellular networks: A two-sided matching game," *IEEE Trans. Mobile Comput.*, vol. 16, no. 10, pp. 2778–2791, Oct. 2017.

[26] B. U. Kazi and G. A. Wainer, "Next generation wireless cellular networks: Ultra-dense multi-tier and multi-cell cooperation perspective," *Wireless Netw.*, vol. 25, no. 4, pp. 2041–2064, May 2019.

[27] D. Gale and L. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[28] F. Kojima and P. A. Pathak, "Incentives and stability in large two-sided matching markets," *Amer. Econ. Rev.*, vol. 99, no. 3, pp. 608–627, May 2009.

[29] Z. Han and K. J. R. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[30] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: Fundamentals and applications," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 52–59, May 2015.

[31] S. G. Glisic, "Wireless networks and matching theory," in *Advanced Wireless Networks*. Hoboken, NJ, USA: Wiley, May 2016, pp. 771–796.

[32] S. Bayat, Y. Li, L. Song, and Z. Han, "Matching theory: Applications in wireless communications," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 103–122, Nov. 2016.

[33] T. Sanguanpuak, S. Guruacharya, N. Rajatheva, M. Bennis, and M. Latva-Aho, "Multi-operator spectrum sharing for small cell networks: A matching game perspective," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3761–3774, Jun. 2017.

[34] B. Di, L. Song, Y. Li, and G. Y. Li, "Non-orthogonal multiple access for high-reliable and low-latency V2X communications in 5G systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2383–2397, Oct. 2017.

[35] T. Hossler, M. Simsek, and G. P. Fettweis, "Matching-based resource allocation for multi-user URLLC in unlicensed frequency bands," in *Proc. 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Oulu, Finland, Aug. 2019, pp. 102–106.

[36] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects, (V9.0.0)*, document 36.814, 3GPP, 2010.

[37] *5G; NR; Physical Layer Measurements, (V15.3.0)*, document 38.215, 3GPP, 2018.

[38] *Technical Specification Group Radio Access Network; Study on New Radio Access Technology; Physical Layer Aspects (Release 14), (V14.2.0)*, document 38.802, 3GPP, 2017.

[39] E. A. Jorswieck, M. Bengtsson, and B. Ottersten, "On the interplay between scheduling, user distribution, csi, and performance measures in cellular downlink," in *Proc. 14th Eur. Signal Process. Conf.*, Florence, Italy, Sep. 2006, pp. 1–5.

[40] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. VTC-Spring. IEEE 51st Veh. Technol. Conf.*, Tokyo, Japan, May 2000, pp. 1085–1089.

[41] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *J. Amer. Stat. Assoc.*, vol. 53, no. 282, pp. 457–481, 1958.

Tom Hößler (Student Member, IEEE) received the Dipl.-Ing. degree in electrical engineering from Technische Universität Dresden (TU Dresden) in 2014. From 2014 to 2015, he worked as a Research Associate with the Fraunhofer Institute for Transportation and Infrastructure Systems IVI. Since 2015, he has been with the Vodafone Chair Mobile Communications Systems, TU Dresden. In 2018, he joined the wireless connectivity group of the Barkhausen Institut, Dresden, Germany. His research interests include dependability theory, ultra-reliable low-latency communications, multi-connectivity, matching theory, and self-organizing networks.

Philipp Schulz (Student Member, IEEE) received the M.Sc. degree in mathematics and the Ph.D. (Dr.-Ing.) degree in electrical engineering from Technische Universität Dresden (TU Dresden), Germany, in 2014 and 2020, respectively. He was a Research Assistant with TU Dresden in the field of numerical mathematics, modeling, and simulation. In 2015, he joined the Vodafone Chair Mobile Communications Systems, TU Dresden, and became a member of the system-level group. His research there focuses on flow-level modeling and the application of queuing theory on communications systems with respect to ultra-reliable low latency communications.

Eduard A. Jorswieck (Fellow, IEEE) was born in Berlin, Germany, in 1975. Since August 2019, he has been the Head of the Chair for Communications Systems and a Full Professor with Technische Universität Braunschweig, Germany. From 2008 until 2019, he was the Head of the Chair of Communications Theory and a Full Professor with Technische Universität Dresden, Germany. His main research interests are in the broad area of communications. He has published more than 120 journal articles, 13 book chapters, 3 monographs, and some 290 conference papers on these topics. He has also been a member of the IEEE SAM Technical Committee since 2015. In 2006, he received the IEEE Signal Processing Society Best Paper Award. Since 2017, he has also been serving as the Editor-in-Chief of the *EURASIP Journal on Wireless Communications and Networking*. He also serves on the editorial board for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Meryem Simsek (Senior Member, IEEE) received the Dipl.-Ing. degree in electrical engineering and information technology and the Ph.D. degree in reinforcement learning-based inter-cell interference coordination in LTE-Advanced heterogeneous networks from the University of Duisburg-Essen, Germany, in 2008 and 2013, respectively. In 2013, she joined Florida International University, USA, as a Post-Doctoral Scientist. Since 2014, she has been a Research Group Leader with the Technische Universität Dresden, Germany. In 2016, she joined the International Computer Science Institute Berkeley as a Senior Research Scientist and UC Berkeley as a Visiting Researcher. Her current research interests include wireless systems, self-organizing networks, and machine learning. she was a recipient of the IEEE Communications Society Fred W. Ellersick Prize in 2015 and the Rising Star in Computer Networking and Communications by N2Women in 2019. She has initiated and is the Chair of the IEEE Tactile Internet Technical Committee and serves as the Vice Chair for the IEEE P1918.1 Standardization Working Group, which she has co-initiated.

Gerhard P. Fettweis (Fellow, IEEE) received the Ph.D. degree under H. Meyr at RWTH Aachen. After a Postdoc at IBM Research, San Jose, he joined TCSI, Berkeley. Since 1994, he has been the Vodafone Chair Professor with TU Dresden. Since 2018, he has also been heads the new Barkhausen Institute. In 2019, he was elected into the DFG Senate. He researches wireless transmission and chip design, and coordinates, e.g., the 5GLab Germany, spun-out 18 startups. He is a member of two German Academies: Academy of Sciences/Leopoldina, Academy of Engineering/acatech.