Wireless Netw (2015) 21:1211–1226
DOI 10.1007/s11276-014-0844-5

Distributed channel assignment for network MIMO: game-theoretic formulation and stochastic learning

Li-Chuan Tseng · Feng-Tsun Chien · Ronald Y. Chang · Wei-Ho Chung · ChingYao Huang · Abdelwaheb Marzouki

Published online: 15 November 2014

Abstract The cooperative frequency reuse among base stations (BSs) can improve the system spectral efficiency by reducing the intercell interference through channel assignment and precoding. This paper presents a game-theoretic study of channel assignment for realizing network multiple-input multiple-output (MIMO) operation under time-varying wireless channel. We propose a new joint precoding scheme that carries enhanced interference mitigation and capacity improvement abilities for network MIMO systems. We formulate the channel assignment problem from a game-theoretic perspective with BSs as the players, and show that our game is an exact potential game given the proposed utility function. A distributed, stochastic learning-based algorithm is proposed where each BS progressively moves toward the Nash equilibrium (NE) strategy based on its own action-reward history only. The convergence properties of the proposed learning algorithm toward an NE point are theoretically and numerically verified for different network topologies. The proposed learning algorithm also demonstrates an improved capacity and fairness performance as compared to other schemes through extensive link-level simulations.

Keywords Network MIMO · Channel selection · Potential games · Stochastic learning

L.-C. Tseng (✉) · F.-T. Chien · C. Huang
Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lctseng@gmail.com

F.-T. Chien
e-mail: ftchien@mail.nctu.edu.tw

C. Huang
e-mail: cyhuang@mail.nctu.edu.tw

R. Y. Chang · W.-H. Chung
Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan
e-mail: rchang@citi.sinica.edu.tw

W.-H. Chung
e-mail: whc@citi.sinica.edu.tw

A. Marzouki
Institut Mines-Télécom, Télécom SudParis, Evry, France
e-mail: abdelwaheb.marzouki@it-sudparis.eu

1 Introduction

Universal frequency reuse is a key technique to improve the throughput of broadband wireless networks. However, frequency reuse among neighboring cells inevitably results in intercell interference (ICI) and degrades the achievable throughput performance. To overcome this problem, ICI management techniques such as ICI coordination and base-station cooperation have been proposed [1, 2]. Base-station cooperation, also known as network multiple-input multiple-output (MIMO), is a multi-antenna signal processing technique that enables several nearby BSs to jointly serve multiple mobile stations (MSs). The implementation of network MIMO may require a partial or full sharing of channel state information (CSI) and data among the BSs.

Much of the research on network MIMO and multicell cooperation has focused on signal processing techniques in an orthogonal frequency-division multiple access (OFDMA) system. The channel assignment for each MS is generally assumed to be determined or treated separately from the network MIMO mechanism. Efficient channel allocation (particularly in a distributed manner) for network MIMO in a multi-antenna multicell environment has not been extensively studied. The aim of this work is

therefore to study the distributed channel allocation problem in network MIMO systems.

The main contributions of this paper are as follows:

- We propose a new joint processing scheme with practical consideration of CSI acquisition where an MS is jointly served by a set of selected BSs. The capacity advantages of the proposed scheme over conventional precoding methods are numerically demonstrated.
- We formulate the channel assignment problem using a game-theoretic approach, and show the existence of Nash equilibrium (NE). Moreover, our proposed utility function induces self-enforcing coordination among players where each player (i.e., the BS) chooses its strategy (i.e., perform channel assignment) independently.
- We develop a stochastic learning (SL)-based algorithm for game-theoretic channel assignment where the players update their strategies simultaneously in each play based on the action-reward history. The convergence behaviors toward NE point are theoretically proven and numerically verified for different network topologies.

The rest of the paper is organized as follows. In Sect. 2, we review related works on precoding in multicell multi-antenna networks as well as those on distributed resource allocation. In Sect. 3, the system model and the proposed joint processing are described. The game-theoretic formulation of the channel allocation problem is presented in Sect. 4 and the SL-based solutions are presented in Sect. 5. Numerical results are provided in Sect. 6. Conclusion is given in Sect. 7.

2 Related works

In network MIMO, the implementation may vary depending on the degree of CSI and data sharing availability. In the static clustering scheme [3], a fixed set of nearby BSs cooperate in jointly serving the users where precoding techniques for single-cell multiuser MIMO systems (e.g., block-diagonalization (BD) [4]) are applied to mitigate the multiuser interference. One disadvantage of static clustering is its requirement of a full sharing of data and CSI within a cluster, which creates a significant overhead on the system operation [5]. The overhead will be even greater if intercluster interference is considered [6].

To reduce the information exchange overhead, partial cooperation has been proposed to avoid the full sharing of CSI and/or data. Kaviani et al. [7] proposed a precoding scheme according to the minimum mean square error (MMSE) criterion and Kerret and Gesbert [8] developed a sparse precoding method which determines the most efficient data sharing patterns, both assuming partial data sharing among the BSs. Distributed MIMO precoding was introduced by Kerret and Gesbert [9] assuming partial CSI sharing but full data sharing. Zakhour et al. [10, 11] proposed a distributed precoding scheme by maximizing the virtual signal-to-interference-and-noise ratio (VSINR) with local CSI. Bjornson et al. [12] developed a network MIMO scheme for large cellular networks, where the precoding vectors are computed in centralized (by a central controller) or fully distributed (by each BS independently) fashion with partial CSI and data.

The realization of network MIMO in an OFDMA system involves an important issue: channel allocation. Traditionally, frequency planning with spatial reuse was considered [13–15] to mitigate the ICI among adjacent cells. Dynamic channel allocation schemes were proposed for cognitive radio networks (CRNs) [16] and network MIMO [12], which requires the presence of a central station for coordinating the sequential strategy updates or negotiations among BSs. The development of self-organized, fully-distributed resource allocation schemes can be facilitated by the application of game theory. Self-organized resource allocation in wireless networks based on reinforcement learning (RL) has been studied [17–23]. Within the RL framework, multiagent Q-learning (MAQL) was applied to CRNs [17] and femtocell networks [18]. MAQL involves the actions of other agents as the external state and thus requires the knowledge of all possible actions of all agents. Also, it suffers from the *curse of dimensionality* since the state space grows exponentially as the number of agents increases, which leads to the decrease in the learning speed and the increase in memory requirements [24]. The stochastic learning (SL), in contrast, adjusts the mixed strategies according to specific update rules, based on the action-reward history. SL has been applied to the game-theoretic study of dynamic spectrum access in CRNs [19, 20] and precoder selection in multiple access channels [21]. The convergence toward pure strategies was shown in [19]. Some update rules have nice convergence properties for specific games. The logit update rule was applied to traffic games in [25] and shown to converge toward perturbed NE. A variation of the procedure in [25] was studied in [26]. In [20, 21], the convergence toward NE point for the update rule proposed in [22] was shown. The learning algorithm in [20, 21] adopts mixed strategy updating rule using normalized reward, and thus requires the knowledge of the maximum of the reward function. An application of SL on both strategy and payoff, referred to as the combined fully-distributed payoff strategy reinforcement learning (CODIPAS-RL), was found for MIMO power loading [23]. Hybrid CODIPAS-RL was applied to heterogeneous 4G networks and the convergence of users' network selection

was observed [27]. While SL algorithms have shown promise for wireless applications in the literature, their applications in distributed resource allocation for multiantenna multicell networks as well as distributed networks of random geometry have not been well studied.

Notations Normal letters represent scalar quantities; upper-case and lower-case boldface letters denote matrices and vectors, respectively. $(\cdot)^T$ and $(\cdot)^H$ stands for the transpose and the conjugate transpose, respectively. $\mathbf{I}$ and $\mathbf{0}$ represent the identity matrix and zero vector with proper size, respectively. $\mathbb{1}_{\{cond\}}$ is the indicator function which equals one if the condition *cond* is satisfied, and zero otherwise.

3 The network MIMO system

3.1 System model

We consider the downlink of an N-cell network MIMO OFDMA system. Each BS is equipped with N_t antennas and each MS is equipped with a single antenna. An MS may be served by multiple BSs and a BS may serve multiple MSs simultaneously in a network MIMO setting. The set of BSs is denoted as $\mathcal{N}$. The time domain is divided into slots and the licensed spectrum is divided into K available orthogonal subchannels each of the same bandwidth. A subchannel may be reused by multiple BSs.

In the network MIMO setting, since BSs and MSs located distant apart cause negligible interference to each other, we consider joint transmission only among nearby BSs to reduce the overhead of data sharing and CSI exchange on the backhaul. For ease of exposition, we make the following definitions:

- Each MS_i feedbacks the CSI to a set of BSs in its coordination set, which is defined as

$$\mathcal{C}_i = \left\{ b \in \mathcal{N} \mid \rho_{ib}^2 \geq \alpha_{th} \rho_{ii}^2 \right\} \tag{1}$$

  where ρ_{ib}^2 is the large-scale channel gain between BS_b and MS_i, which can be obtained by averaging over the estimated channel gain at the receiver, and the threshold $0 < \alpha_{th} \leq 1$ is a system design parameter.
- Each MS_i receives the data from its service set, which is defined as

$$\mathcal{D}_i = \left\{ b \in \mathcal{N} \mid \rho_{ib}^2 \geq \beta_{th} \rho_{ii}^2 \right\} \tag{2}$$

  where $\beta_{th} \geq \alpha_{th}$ and $\mathcal{D}_i \subseteq \mathcal{C}_i$.

In the network MIMO system, a BS_b ($b \in \mathcal{C}_i$) can mitigate the interference to the MSs in the coverage area of the other BSs in $\mathcal{C}_i$ through proper precoder designs. An illustrative example of the network MIMO system with joint processing is given in Fig. 1.

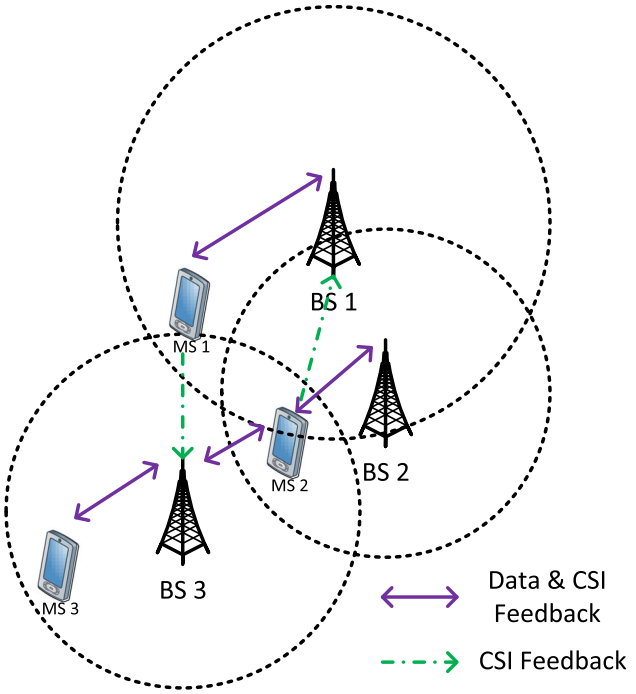


Fig. 1 Illustration of distributed channel assignment with joint precoding in multicell networks. For $\text{MS}_1, \mathcal{C}_1 = \{1,3\}$ and $\mathcal{D}_1 = \{1\}$, where BS_1 and BS_3 both receive CSI feedback from MS_1 and perform interference mitigation but only BS_1 serves MS_1. For $\text{MS}_2, \mathcal{C}_2 = \{1,2,3\}$ and $\mathcal{D}_2 = \{2,3\}$, where BS_2 and BS_3 jointly serve MS_2 while all three BSs perform interference mitigation. For $\text{MS}_3, \mathcal{C}_3 = \{3\}$ and $\mathcal{D}_3 = \{3\}$, where only BS_3 serves MS_3

To reflect a practical wireless network, our system model incorporates the following considerations:

1. The channel state is time-varying so that the channel condition may change before the channel selection process is accomplished. We consider that the coherence bandwidth is greater than the total bandwidth of subchannels available for selection (i.e., a frequency-flat fading channel) for notational and modeling simplicity. Note, however, that the proposed framework is applicable to systems with any coherence bandwidth or coherence time conditions.
2. The number of cells, N, is unknown.
3. Each BS selects the channel independently and simultaneously, in contrast to a coordinated joint decision or sequential updates.

3.2 Transmitter precoding

In the network MIMO system considered in [11], only data are shared among the BSs and the precoding vector is calculated separately at each BS. Here, we propose a joint

processing method in which the BSs in the serving set of each user exchange their knowledge of CSI and determine the precoding vector jointly. Similar to [11], a power splitting procedure is considered which allows each BS to split its transmission power among the MSs that it needs to serve. Let P_b be the transmission power of BS_b on one subchannel. We adopt a simple equal-power splitting method so that the power allocated to MS_i by BS_b is given by

$$P_{ib} = \frac{P_b}{\sum_{i=1}^{N} \mathbb{1}_{\{\mathcal{D}_i \ni b\}}}, \quad \forall i \text{ s.t. } \mathcal{D}_i \ni b. \tag{3}$$

Signal transmission in the multicell network MIMO system is modeled as follows. Let $\mathbf{h}_{\mathbf{ib}}^{\mathbf{k}} \in \mathbb{C}^{\mathbf{N_t \times 1}}$ represent the channel from BS_b to MS_i on subchannel k. The symbol x_i denotes the data intended for MS_i, where $\mathbb{E}[|x_i|^2] = 1$ and $\mathbb{E}[x_i^* x_j] = 0, \forall i \neq j$. The data symbol x_i is precoded by precoders $\mathbf{w_{ib}} \in \mathbb{C}^{\mathbf{N_t \times 1}}, \forall \mathbf{b} \in \mathcal{D}_\mathbf{i}$. Let $D_j = |\mathcal{D}_j|$ be the cardinality of the serving set of MS_j. Then, the collective channel from the BSs in $\mathcal{D}_j$ to MS_i on subchannel k can be expressed as

$$\mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{j}}^{\mathbf{k}} = \left[\sqrt{\mathbf{P_{jb_1}}} \left(\mathbf{h}_{\mathbf{ib_1}}^{\mathbf{k}} \right)^{\mathbf{T}}, \ldots, \sqrt{\mathbf{P_{jb_{D_j}}}} \left(\mathbf{h}_{\mathbf{ib_{D_j}}}^{\mathbf{k}} \right)^{\mathbf{T}} \right]^{\mathbf{T}} \tag{4}$$

and the collective precoding vector for MS_i is

$$\mathbf{w_i} = \left[\mathbf{w}_{\mathbf{ib_1}}^{\mathbf{T}}, \ldots, \mathbf{w}_{\mathbf{ib_{D_i}}}^{\mathbf{T}} \right]^{\mathbf{T}}. \tag{5}$$

Let $a_i(n)$ be the selected channel for MS_i (i.e., the action taken by BS_i) at slot n. For notational brevity, we will hereafter discard the timing dependence of the action $a_i(n)$ in occasions without ambiguity. Also, in consideration of a frequency-flat fading channel mentioned previously for notational and modeling simplicity without compromising the generality of the proposed framework, we will discard the subchannel index k. The discrete-time baseband signal received by MS_i is given by

$$y_i = \mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{i}}^{\mathbf{T}} \mathbf{w_i} \mathbf{x_i} + \sum_{\mathbf{j=1, j \neq i}}^{\mathbf{N}} \mathbb{1}_{\{\mathbf{a_i = a_j}\}} \mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{j}}^{\mathbf{T}} \mathbf{w_j} \mathbf{x_j} + \mathbf{z_i} \tag{6}$$

where the first term is the desired signal, the second term represents the ICI, and z_i is additive complex Gaussian noise with variance σ^2. Therefore, the signal-to-interference-and-noise ratio (SINR) at MS_i can be formulated as

$$\gamma_i = \frac{\|\mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{i}}^{\mathbf{T}} \mathbf{w_i}\|^2}{\sum_{j=1, j \neq i}^{N} \mathbb{1}_{\{a_i = a_j\}} \|\mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{j}}^{\mathbf{T}} \mathbf{w_j}\|^2 + \sigma^2}. \tag{7}$$

The achievable capacity for MS_i in bits/s/Hz is given by

$$R_i = \log_2 \left(1 + \frac{\gamma_i}{\Gamma} \right) \tag{8}$$

where $\Gamma = \ln(5\text{BER})/1.5$ is a function of the required bit error rate (BER), often known as the SINR gap [28].

We denote the precoding vector $\mathbf{w}_i$ for MS_i by $\mathbf{w}_i = \mu_i \hat{\mathbf{w}}_i$, where μ_i is an adjustment factor to maintain the per-BS power constraint and $\hat{\mathbf{w}}_i$ is the unit-norm vector that maximizes the *modified signal-to-leakage-and-noise ratio* (mSLNR). Different from the SLNR in [29], we consider the mSLNR to reflect a practical network MIMO operation, which is defined in terms of the signal power received by MS_i and the *available* information to BS_i about the interference caused to other MSs (produced by the signals from $\mathcal{D}_i$ intended for MS_i) plus the noise power. The distinction on the interference part is made to reflect the fact that not all CSI can be acquired by the BSs in $\mathcal{D}_i$ and thus the interference powers imposed on other users may not be available. Specifically, in our consideration a BS in $\mathcal{D}_i$ can acquire the CSI to MS_j ($i \neq j$) only if this BS is also in $\mathcal{C}_j$. Mathematically, $\hat{\mathbf{w}}_i$ is given by

$$\hat{\mathbf{w}}_i = \underset{\|\mathbf{w}\|=\mathbf{1}}{\operatorname{argmax}} \underbrace{\frac{\left\| \mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{i}}^{\mathbf{T}} \mathbf{w} \right\|^2}{\sigma^2 + \sum_{j=1, j \neq i}^{N} \mathbb{1}_{\{a_i = a_j\}} \left\| \tilde{\mathbf{h}}_{j,\mathcal{D}_i}^{T} \mathbf{w} \right\|^2}}_{\text{mSLNR of MS}_\text{i}} \tag{9}$$

where

$$\tilde{\mathbf{h}}_{j,\mathcal{D}_i} = \left[\sqrt{P_{ib_1}} \breve{\mathbf{h}}_{jb_1}^{T}, \ldots, \sqrt{P_{ib_{D_i}}} \breve{\mathbf{h}}_{jb_{D_i}}^{T} \right]^{T} \tag{10}$$

with

$$\breve{\mathbf{h}}_{jb} = \begin{cases} \mathbf{h_{jb}}, & \text{if } \mathrm{b} \in \mathcal{C}_\mathrm{i} \cap \mathcal{C}_\mathrm{j}, \\ \mathbf{0}, & \text{otherwise.} \end{cases} \tag{11}$$

The vector $\breve{\mathbf{h}}_{jb}$ reflects our mSLNR consideration; that is, it is equal to the CSI when this information can be collected (via feedbacks or backhaul communications), and zero otherwise.

The solution to (9) is given by

$$\hat{\mathbf{w}}_i = \frac{\mathbf{K_i^{-1}} \mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{i}}}{\left\| \mathbf{K_i^{-1}} \mathbf{h}_{\mathbf{i},\mathcal{D}_\mathbf{i}} \right\|} \tag{12}$$

where $\mathbf{K_i} = \sigma^2 \mathbf{I} + \sum_{\mathbf{j \neq i}} \mathbb{1}_{\{\mathbf{a_i = a_j}\}} \mathbf{h}_{\mathbf{j},\mathcal{D}_\mathbf{i}} \mathbf{h}_{\mathbf{j},\mathcal{D}_\mathbf{i}}^{\mathbf{H}}$. We then employ a heuristic approach similar to [6] to obtain the adjustment factor μ_i as

$$\mu_i = \frac{1}{\max\{\|\mathbf{w}_{\mathbf{ib_1}}\|, \|\mathbf{w}_{\mathbf{ib_2}}\|, \ldots, \|\mathbf{w}_{\mathbf{ib_{D_i}}}\|\}}. \tag{13}$$

Note that the multicell precoding scenario considered in [10] is a special case of our proposed method. In this local precoding scheme, each BS's knowledge of CSI is limited to the channel between itself and the MSs under its coverage. Each BS's CSI is obtained through a feedback mechanism and maintained locally. By setting $\beta_{th} > 1$ in

Table 1 Summary of notations in game-theoretic formulation

Symbol	Meaning
$\mathcal{H}$	External state (channel state) space
$\mathbf{H}$	Random matrix for the channel state
$\mathcal{N}$	Set of players
$\mathcal{A}_i$	Set of actions of player i
$s_i \in \mathcal{A}_i$	An element of $\mathcal{A}_i$
$a_i(n) \in \mathcal{A}_i$	Action (channel assignment) of player i at slot n
$a_{-i}(n) \in \mathcal{A}_i$	Actions of players except for i at slot n
$\mathcal{P}_i := \Delta(\mathcal{A}_i)$	Set of probability distribution over $\mathcal{A}_i$
$\mathbf{p_i(n)} \in \mathcal{P}_\mathbf{i}$	Mixed strategy of player i at slot n
$r_i(n) \in \mathbb{R}$	Instantaneous reward of player i at slot n
$\hat{\mathbf{u}}_i(n) \in \mathbb{R}^{\lvert\mathcal{A}_i\rvert}$	Estimated utility vector of player i at slot n
(ϵ_i, η_i)	Learning rates of player i

our system, the serving set of each MS will consist of its home BS only and thus the system reduces to local precoding. The performance of local precoding may be limited since the neighboring BSs of an MS act only as a source of interference without providing any useful data streams. The performance comparison of local precoding and joint processing is presented in Sect. 6.

4 Channel assignment for network MIMO

In this section, we present the game-theoretic formulation of the self-organized channel assignment to realize the network MIMO scheme described in Sect. 3. Our objective is to devise a distributed channel assignment strategy that takes into account the effect of ICI. We summarize our notations related to the game formulation in Table 1.

4.1 Game-theoretic formulation

We model the channel assignment as a noncooperative game with external state, expressed as a 4-tuple:

$$\mathcal{G} = \left(\mathcal{H}, \mathcal{N}, \{\mathcal{A}_\rangle\}_{\rangle\in\mathcal{N}}, \{\sqcap_\rangle\}_{\rangle\in\mathcal{N}}\right)$$

where $\mathcal{H}$ is the external state (channel state) space, $\mathcal{N} = \{1, \ldots, N\}$ is the set of players (BSs), $\mathcal{A}_i = \{1, \ldots, K\}$ is the set of actions (selections of channels) that player i can take, and u_i is the ergodic utility function of player i defined as the expected reward over the time-varying channel state, i.e.,

$$u_i(a_i, a_{-i}) \triangleq \mathbb{E}_\mathbf{H}[r_i(a_i, a_{-i}; \mathbf{H})] \tag{14}$$

where a_{-i} represents the actions of other players except for i, and $r_i : \times_{i\in\mathcal{N}}\mathcal{A}_i \mapsto \mathbb{R}$ represents the instantaneous reward function for player i under a given channel state $\mathbf{H}$. Note that (14) does not require any specification of slow/fast fading or frequency-flat/selective fading conditions of the channel. Intuitively, the achievable capacity in (8) may be considered as the reward function. However, we notice that in [19] the interference terms related to the action of player i are treated as the cost of player i, and the negation of summed cost is defined as the reward. The advantage of this reward function design lies in that, during the learning procedure, in addition to maximizing its own rate, a player now also tends to minimize the interference generated to other players due to its action. Therefore, implicit *coordination* can be achieved even with a noncooperative game formulation. In this paper, with the joint processing scenario and inspiration by [19], we propose to design the reward function as

$$r_i(a_i, a_{-i}; \mathbf{H}) \triangleq -\left[\sum_{\mathbf{j}=\mathbf{1},\mathbf{j}\neq\mathbf{i}}^{\mathbf{N}} \mathbf{I_{j\to i}} + \sum_{\substack{\mathbf{j}=\mathbf{1},\\ \mathcal{D}_\mathbf{j}\cap\mathcal{C}_\mathbf{i}\neq\emptyset}}^{\mathbf{N}} \sum_{\mathbf{m}=\mathbf{1},\ \mathbf{m}\neq\mathbf{i},\mathbf{j}}^{\mathbf{N}} \mathbf{I_{j\to m}}\right] \tag{15}$$

where

$$I_{j\to i} \triangleq \mathbb{1}_{\{a_i=a_j\}} \frac{\left\|\tilde{\mathbf{h}}_{i,\mathcal{D}_j}^T \mathbf{w_j}\right\|^2}{\left\|\tilde{\mathbf{h}}_{j,\mathcal{D}_j}^T \mathbf{w_j}\right\|^2} \tag{16}$$

is the interference caused at MS_i by the signal intended for MS_j normalized by the received signal power of MS_j. The considered reward function is composed of the I-values that may vary when player i changes its action. The first term in (15) accounts for the total interference caused at MS_i as a result of external BSs. In our design, the first term represents the *selfish* motivation to minimize the sum of incoming interference, which aligns closely with the capacity function (8), i.e., the lower interference leads to higher capacity. On the other hand, the second term in (15) is the *altruistic* part of the reward function, which accounts for the interference imposed on other MSs by the signal intended for MS_j when $\mathcal{D}_j \cap \mathcal{C}_i \neq \emptyset$. Note that this altruistic term varies with player i's action a_i, since the precoder design for a link j such that $\mathcal{D}_j \cap \mathcal{C}_i \neq \emptyset$ is changed whenever a different a_i is adopted, as revealed in (9). The two terms in (15) together characterize the overall effects of interference due to the action a_i. Consequently, maximizing the reward function will lead to an assignment of subchannels that causes minimum interference impacts.

4.2 Analysis of Nash equilibrium

We assume that the players (i.e., the BSs) in the proposed game are selfish and rational. In other words, they will

compete to maximize their individual utilities, i.e., maximizing their own throughput while reducing the interference generated to others.

Definition 1 An action profile $\mathbf{a}^* = (\mathbf{a}_1^*, \ldots, \mathbf{a}_N^*)$ is a pure strategy Nash equilibrium (NE) point of the noncooperative game $\mathcal{G}$ if and only if no player can improve its utility by deviating unilaterally, i.e.,

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*), \quad \forall i \in \mathcal{N}, \forall a_i \in \mathcal{A}_i \setminus \{a_i^*\}. \tag{17}$$

With the reward function defined in (15), we show the existence of an NE point for the proposed game in the following proposition.

Proposition 1 The proposed channel selection game $\mathcal{G}$ is an exact potential game (EPG) with at least one pure strategy NE point.

Proof For a channel assignment profile (a_i, a_{-i}), consider the following function $\Phi : \times_{i\in\mathcal{N}}\mathcal{A}_i \mapsto \mathbb{R}$ for the game $\mathcal{G}$:

$$\Phi(a_i, a_{-i}) = \mathbb{E}_{\mathbf{H}}\left[-\sum_{j=1}^{N}\sum_{m=1, m\neq j}^{N} I_{j\to m}\right]. \tag{18}$$

Observing that player i's change does not affect the precoder of MS_j if $\mathcal{D}_j \cap \mathcal{C}_i = \emptyset$, we define

$$r_{-i}(a_{-i}; \mathbf{H}) \triangleq -\sum_{\substack{\mathbf{j}=\mathbf{1}, \\ \mathcal{D}_\mathbf{j}\cap\mathcal{C}_\mathbf{i}=\emptyset}}^{\mathbf{N}} \sum_{\substack{\mathbf{m}=\mathbf{1}, \\ \mathbf{m}\neq\mathbf{i},\mathbf{j}}}^{\mathbf{N}} \mathbf{I}_{\mathbf{j}\to\mathbf{m}}. \tag{19}$$

Considering a unilateral strategy for player i that changes its action unilaterally from a_i to $\breve{a}_i$, we have

$$\begin{aligned} &u_i(\breve{a}_i, a_{-i}) - u_i(a_i, a_{-i}) \\ &= \mathbb{E}_{\mathbf{H}}[r_i(\breve{a}_i, a_{-i}; \mathbf{H})] - \mathbb{E}_{\mathbf{H}}[\mathbf{r_i}(\mathbf{a_i}, \mathbf{a_{-i}}; \mathbf{H})] \\ &= \mathbb{E}_{\mathbf{H}}[r_i(\breve{a}_i, a_{-i}; \mathbf{H}) + \mathbf{r_{-i}}(\mathbf{a_{-i}}; \mathbf{H})] \\ &\quad - \mathbb{E}_{\mathbf{H}}[r_i(a_i, a_{-i}; \mathbf{H}) + \mathbf{r_{-i}}(\mathbf{a_{-i}}; \mathbf{H})] \\ &= \Phi(\breve{a}_i, a_{-i}) - \Phi(a_i, a_{-i}). \end{aligned} \tag{20}$$

From (20) and by the definition of an EPG [30], $\mathcal{G}$ is an EPG with Φ as its potential function and the existence of a pure strategy NE point is guaranteed. □

One important property of a potential game is that the interests of players align to a global objective: maximization of the potential function. For example, with (18), the players in $\mathcal{G}$ actually minimize the total cost in the system. This property suggests the possibility of distributed learning toward the equilibrium. Note that a mixed strategy NE, which is the type of equilibrium the learning algorithm computes, always exists for a noncooperative finite game. Furthermore, the convergence toward a pure strategy NE is observed through numerical simulations, as to be presented in Sect. 6.

We can easily extend the proposed channel assignment game into a mixed strategy form as in [22]. Let $\mathbf{p_i}(\mathbf{n}) = [\mathbf{p_{i,1}}(\mathbf{n}), \ldots, \mathbf{p_{i,K}}(\mathbf{n})]^{\mathbf{T}}, \forall \mathbf{i} \in \mathcal{N}$ be the channel assignment probability vector for player i, where $p_{i,s_i}(n)$ is the probability that player i selects strategy $s_i \in \mathcal{A}_i$ at slot n. Then, the mixed extension of utility function is defined on $\times_{i\in\mathcal{N}}\mathcal{P}_i$, where $\mathcal{P}_i$ is the set of probability distribution over the action space of player i. Let $\mathbf{P}(\mathbf{n}) = [\mathbf{p_1}(\mathbf{n}), \ldots, \mathbf{p_N}(\mathbf{n})]$ be the mixed strategy profile of $\mathcal{G}$, we denote the mixed extension of utility by $\psi_i(\mathbf{P})$, i.e.,

$$\psi_i(\mathbf{P}) = \sum_{\mathbf{a_1},\ldots,\mathbf{a_N}} \mathbf{u_i}(\mathbf{a_1}, \ldots, \mathbf{a_N}) \prod_{\mathbf{j}=\mathbf{1}}^{\mathbf{N}} \mathbf{p_{j,a_j}}. \tag{21}$$

Let $\mathbf{P_{-i}}$ be the mixed strategy of players except for player i, we have the definition of NE in mixed strategy as follows.

Definition 2 A strategy profile $\mathbf{P}^*$ is a mixed-strategy Nash equilibrium (NE) point of the noncooperative game $\mathcal{G}$ if and only if

$$\psi_i(\mathbf{p_i^*}, \mathbf{P_{-i}^*}) \geq \psi_i(\mathbf{p_i}, \mathbf{P_{-i}^*}), \quad \forall \mathbf{i} \in \mathcal{N}, \forall \mathbf{p_i} \in \mathcal{P}_\mathbf{i} \setminus \{\mathbf{p_i^*}\}. \tag{22}$$

Later in Sect. 5, a mechanism is studied to reach a Nash equilibrium of the game.

4.3 Acquisition of the interference information

It is practically difficult to obtain the exact value of the reward function in (15) for each player, as the calculation of (15) relies on complete knowledge of CSI while the CSI feedback is limited to only BSs in the coordination set. In consideration of CSI availability and the geographic relationship of two MSs, it is useful to consider the following approximation for practical implementation:

$$I_{j\to i} \approx 0, \quad \forall j \in \mathcal{N} \text{ s.t. } \mathcal{D}_\mathrm{j} \cap \mathcal{C}_\mathrm{i} = \emptyset. \tag{23}$$

Note that $\mathcal{D}_j \cap \mathcal{C}_i = \emptyset$ largely indicates that MS_i and MS_j are geographically farther apart. Then, by combining the two terms in (15), the instantaneous reward function in (15) can be approximated by[1]

$$r_i \approx -\sum_{j=1, \mathcal{D}_j\cap\mathcal{C}_i\neq\emptyset}^{N} I_j^{out} \tag{24}$$

where

[1] Note that with the approximated utility function the existence of a pure strategy NE is no longer guaranteed theoretically. However, as will be shown in Sect. 6, convergence to NE is observed numerically.

$$I_j^{out} = \sum_{m=1, m\neq j}^{N} I_{j\to m}. \tag{25}$$

The expression in (25) defines the (normalized) outward interference of player j. In other words, the reward function of player i takes into account the players whose coordination set overlaps with the service set of player i. A two-step protocol can therefore be established:

1. Each player j calculates its own I_j^{out} based on the CSI feedback, and
2. Each player exchanges the information with other players.

5 Stochastic learning-based channel assignment algorithm

There has been much interest in designing learning algorithms toward NE in noncooperative games. However, the external state (CSI) is unknown and the action is selected by each player simultaneously and independently in each play. Therefore, previous algorithms requiring complete information and implicit ordering of acting players (e.g., those based on better response dynamics [30] and fictitious play [31]) may not be feasible in our self-organized multicell resource allocation problem. In this section, we develop a distributed SL-based algorithm where the BSs move toward the equilibrium strategy profile based on their individual action-reward history. Our algorithm adopts a two-time-scale stochastic approximation [32] with the utility update step based on [33] and the action selection scheme based on [34]. For further discussions of the SL-based algorithm related to this work, the reader is referred to [26, 35–37], as well as [38] for a comprehensive literature review.

5.1 Algorithm description

The proposed SL-based channel assignment algorithm is described in Algorithm 1. In each play, the channel is selected based on the probability distribution over the set of channels. At the completion of each play, a player obtains the instantaneous reward and updates the estimated utility vector $\hat{\mathbf{u}}_i(n)$ as well as the channel assignment probability vector $\mathbf{p_i}(\mathbf{n}+\mathbf{1})$ for the next play, according to the update rules specified in (26). A straightforward interpretation of the rules is that the utility estimation serves as a reinforcement signal so that higher utility (lower cost) leads to higher probability in the next play. Notably, the proposed learning algorithm is distributed: the channel assignment is done by each player based on the individual action-reward experience, instead of a joint decision. Moreover, although the SL-based algorithm proposed in [22] also converges to NE points for potential games, its probability update rule requires the normalization of the instant reward such that its value will lie in $[0, 1]$. This requirement of normalization makes the algorithm inapplicable when the extreme values of reward functions are unavailable. This restriction however does not apply to the proposed algorithm due to a different probability update rule.

Algorithm 1 Stochastic Learning toward NE

1: Initially, set $n = 0$. Set the channel assignment probability vector and utility estimation as

$$p_{i,s_i}(0) = 1/|\mathcal{A}_i|, \hat{u}_{i,s_i}(-1) = 0, \forall i \in \mathcal{N}, s_i \in \mathcal{A}_i.$$

2: At the beginning of the nth slot, each player selects an action $a_i(n)$ according to the current channel assignment probability $\mathbf{p}_i(n)$.

3: In each slot, each BS transmits data. At the end of each slot, each BS receives the instantaneous reward $r_i(n)$ specified by (15) depending on the precoding scheme.

4: All BSs update their channel assignment probability vector and utility estimation according to the rules:

$$\begin{cases} \hat{u}_{i,s_i}(n) - \hat{u}_{i,s_i}(n-1) = \eta_i \mathbb{1}_{\{a_i(n)=s_i\}} \left(r_i(n) - \hat{u}_{i,s_i}(n-1)\right) \\ p_{i,s_i}(n+1) = \dfrac{p_{i,s_i}(n)(1-\epsilon_i)^{-\hat{u}_{i,s_i}(n)}}{\sum_{s_i' \in \mathcal{A}_i} p_{i,s_i'}(n)(1-\epsilon_i)^{-\hat{u}_{i,s_i'}(n)}} \end{cases} \tag{26}$$

where ϵ_i and η_i are the learning rates for action probability and utility estimation, respectively.

Note that the complexity of the proposed SL-based channel assignment algorithm is dominated by the computation of the precoding vector to find the reward function (Step 3 of Algorithm 1) for a number of iterations before convergence (Step 4 of Algorithm 1). The expected convergence time is roughly proportional to the initial value of the potential function and inversely proportional to the learning rates [39]. The convergence time also depends on other factors such as the number of players, since the Lyapunov function (i.e., the negative potential function) of the potential game is the sum of normalized interference over all players.

5.2 Convergence properties of the proposed algorithm

Convergence toward NE points is an important feature of the proposed learning algorithm. Similar to the discussions in [22] and [20], here we theoretically demonstrate the convergence properties of the proposed SL-based algorithm. First, by using the ordinary differential equation (ODE) approximation we characterize the long-term behavior of the sequence $\{\mathbf{P}(n)\}$. Second, we establish a sufficient condition for the arrival at NE points for the proposed learning algorithm and prove that the game $\mathcal{G}$ satisfies this condition.

Proposition 2 With sufficiently small ϵ_i and η_i, the piecewise linearly interpolated process of the sequence $p_{i,s_i}(n)$ is bounded with high probability within arbitrarily small vicinity of the flow induced by the following ODE:

$$\frac{dp_{i,s_i}(t)}{dt} = p_{i,s_i}(t)[\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \psi_{\mathbf{i}}(\mathbf{P})] \tag{27}$$

where $\mathbf{e}_{\mathbf{s_i}}$ is a unit probability vector (of appropriate dimension) with the s_i-th component being unity and all others zero. The initial condition is given by $\boldsymbol{P}(0) = \boldsymbol{P}_0$, where $\boldsymbol{P}_0$ is the initial channel assignment probability matrix.

Proof See [40], Sect. 4.3]. □

Note that the ODE in (27) is the *replicator equation* [38] in which the probability of taking one strategy grows if this strategy's current estimated utility is larger than the average utility over all strategies and declines otherwise. Compared to the best response dynamics [30] where a player changes its strategy in the next iteration to the best action according to other players' action, with the replicator dynamics, a player selects an action according to a probability distribution over the action set, and adjusts the weighting for each possible action in each iteration based on the estimated utility.

Proposition 3 The replicator dynamics have the following properties: [38]

1. All Nash equilibria are stationary points;
2. All (Lyapunov) stable stationary points are Nash equilibria. More generally, any stationary point that is the limit of a path that originates in the interior is a Nash equilibrium.

Proposition 3 is an instance of the Folk theorems in evolutionary game theory [38], and these properties follow directly from the replicator equation in (27). For an intuitive explanation, observe that $\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}})$ is the expected reward function of player i if it employs pure strategy s_i while other player $j, \forall j \in \mathcal{N}, j \neq i$ employs a mixed strategy $\mathbf{p_j}$. From the definition of Nash equilibrium, the condition

$$\psi_i(\mathbf{e}_{\mathbf{s_i^*}}, \mathbf{P}^*_{-\mathbf{i}}) = \psi_{\mathbf{i}}(\mathbf{P}^*), \quad \forall \mathbf{i} \in \mathcal{N}, \mathbf{s_i} \in \mathcal{A}_\mathbf{i} \text{with} p^*_{\mathbf{i},\mathbf{s_i}} > 0 \tag{28}$$

must hold for an NE strategy profile $\mathbf{P}^*$. Therefore any Nash equilibrium must lead the right-hand side of (27) to zero, and thus constitutes a stationary point of (27). It is worth noting that, for a mixed-strategy NE, all survived pure strategies (i.e. s_i with $p_{i,s_i} > 0$) of player i perform equally well when other players follow the mixed strategy $\mathbf{P}^*_{-\mathbf{i}}$.

Proposition 2 investigates the convergence behavior of the discrete-time learning algorithm toward the trajectory of the replicator dynamics (27), and Proposition 3 states the relation between the stationary point of the trajectory and NE. Next, we study the sufficient condition for the convergence of the learning algorithm toward NE in the following two propositions.

Proposition 4 Suppose that there exists a bounded differentiable function $\Psi : \mathbb{R}^{|\mathcal{A}|} \to \mathbb{R}$ such that

$$\Psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) = \frac{\partial \Psi(\mathbf{P})}{\partial \mathbf{p}_{\mathbf{i},\mathbf{s_i}}}, \quad \forall \mathbf{i} \tag{29}$$

is positively correlated with $\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}})$ in the sense that $\psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \psi(\mathbf{e}_{\mathbf{s'_i}}, \mathbf{P}_{-\mathbf{i}}) > \mathbf{0}$ if and only if $\Psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \Psi(\mathbf{e}_{\mathbf{s'_i}}, \mathbf{P}_{-\mathbf{i}}) > \mathbf{0}$. Then, the SL-based algorithm converges weakly to an NE point of a noncooperative game.

Proof See the Appendix. □

Proposition 4 establishes a sufficient condition that guarantees the convergence toward NE. In what follows, we prove that the proposed channel assignment game $\mathcal{G}$ satisfies this condition and hence it converges weakly to an NE point by using the SL-based channel assignment algorithm.

Proposition 5 When applied to EPGs, the proposed SL-based channel assignment algorithm converges weakly to an NE point.

Proof For EPGs, let $\Psi(\mathbf{P})$ be the mixed extension of the potential function,

$$\Psi(\mathbf{P}) = \sum_{\mathbf{a_1},\ldots,\mathbf{a_N}} \Phi(\mathbf{a_1}, \ldots, \mathbf{a_N}) \prod_{\mathbf{j}=\mathbf{1}}^{\mathbf{N}} \mathbf{p}_{\mathbf{j},\mathbf{a_j}}. \tag{30}$$

From (20), we have

$$\Psi(\mathbf{e}_{\mathbf{s}_i'}, \mathbf{P}_{-\mathbf{i}}) - \Psi(\mathbf{e}_{\mathbf{s}_i}, \mathbf{P}_{-\mathbf{i}}) = \psi_{\mathbf{i}}(\mathbf{e}_{\mathbf{s}_i'}, \mathbf{P}_{-\mathbf{i}}) - \psi_{\mathbf{i}}(\mathbf{e}_{\mathbf{s}_i}, \mathbf{P}_{-\mathbf{i}}),$$
$$\forall i \in \mathcal{N}, s_i, s_i' \in \mathcal{A}_i, \quad (31)$$

which satisfies the condition in Proposition 4 and completes the proof. □

Remarks

1. The weak convergence in Proposition 4 is in the sense of convergence in law (i.e., convergence in distribution) [41, p. 329].
2. Since the ordinal potential game (OPG) [30] satisfies the condition in Proposition 4, the proposed learning procedure can be applied to problems formulated as OPG, not just EPG.
3. Propositions 4 and 5 coupled with the Folk theorem for multi-population games [38] guarantee convergence toward a *pure-strategy* NE.
4. The learning rates (step sizes) (ϵ_i, η_i) play an important role in the convergence behavior of the SL-based learning algorithm. In particular, smaller step sizes lead to a slower convergence. The choice of learning rates poses a trade-off between accuracy and speed, and may be determined by training in practice.
5. While the stochastic learning process with decreasing step sizes will converge with probability one, a constant step size is useful and often preferable in engineering applications to achieve faster convergence [42]. The stochastic approximation with constant step size does not guarantee that the linearly interpolated process is an asymptotic pseudo-trajectory , a notion introduced by Benaïm and Hirsch [43] for analyzing the long term behavior of stochastic approximation processes with decreasing step size, of the flow induced by the ODE (27). However, the considered stochastic learning algorithm with a small constant step size still admits weak convergence in the sense that with probability close to 1, as the step size approaches zero, the linearly interpolated process of the update rule for the channel selection probability will track the trajectories of the ODE (i.e., the replicator dynamics equation) with an error bounded above by some arbitrarily small fixed positive real value. More details can be found in [22, Remark 3.1] [44, Theorem 2.3] [45].

6 Numerical results and discussions

In this section, our theoretical developments are numerically verified in hexagonal cellular networks as well as distributed networks of random geometry. Universal frequency reuse is adopted in our link-level simulations. The simulation setup follows the 3GPP model [46] and is summarized in Table 2.

Table 2 The simulation setup

Cellular parameters	
Number of cells, N	19 (wrap-around)
Cell radius, R_{BS}	500 m
Min. MS to home BS distance	$0.7R_{BS}$
Number of Tx antennas, N_t	2
OFDMA parameters	
FFT size	128
Carrier frequency	2 GHz
Subcarrier spacing	15 kHz
Number of subchannels	6
Number of subcarriers per subch.	12
Subch. for network MIMO mode	Subch. 1 & 2 ($K = 2$)
Channel model parameters	
PathLoss (dB)	$34.5 + 35 \log_{10} d$ (d in m)
Shadowing SD	8 dB
Speed of MSs	3 km/h
Fast fading	Ray-based model (Sect. 5 of [46])
Power control parameters	
Trans. power	46 dBm
Thermal noise power	−174 dBm/Hz
Other parameters	
Thresholds for coordination	$\alpha_{th} = 0.1$, $\beta_{th} = 0.3$ (default)
Learning rates	$\epsilon_i = \eta_i = 0.1, \forall i \in \mathcal{N}$

6.1 Convergence behaviors of the proposed learning algorithm

We plot the evolution of the channel assignment probability (i.e., the mixed strategies) of the proposed stochastic learning algorithm for four arbitrarily selected players in Fig. 2. It is observed that, with equal initial probabilities, the channel assignment probabilities converge to a pure strategy in around 200–300 iterations. For other players in the game which are not shown, a similar convergence result is also observed. Note that the learning stage is generally a minor and manageable overhead inherent to any learning-based algorithm, as the time required for convergence is typically a small fraction of the total operation time.

Figure 3 shows the evolution of the estimated cost vector (i.e., $-\hat{\mathbf{u}}_i$) of two selected players. As can be seen, the BSs tend to select the channel with lower estimated cost (solid lines). Figures 2 and 3 demonstrate that, with high probability, mutually interfering cells can coordinate their transmissions on different channels even without negotiations.

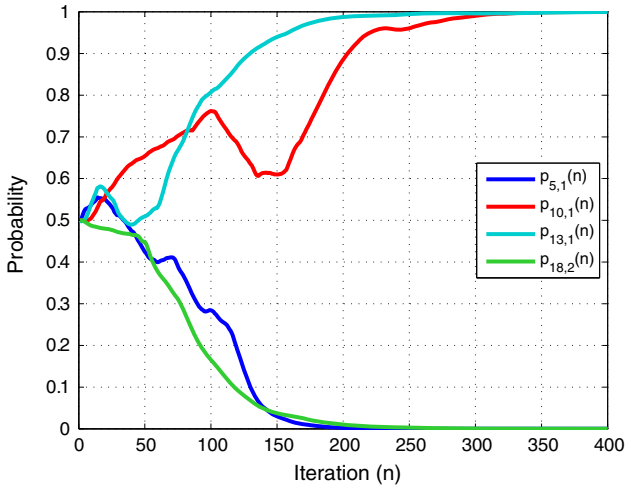


Fig. 2 Evolution of the mixed strategies (probability of taking different actions) of four selected players when joint processing is adopted

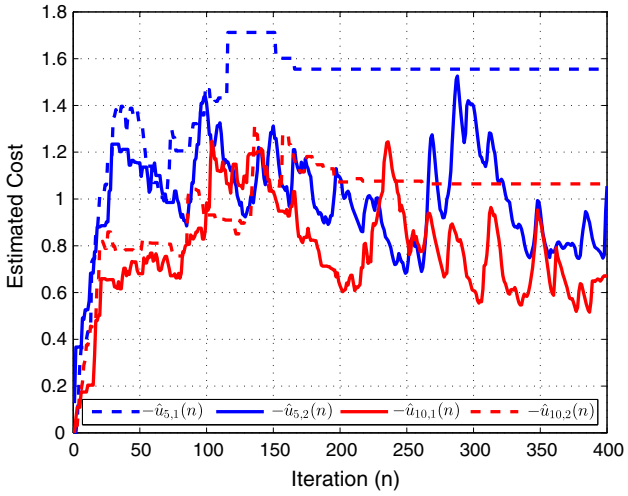


Fig. 3 Evolution of the estimated cost of taking different actions for two selected players (marked by *blue* and *red colors*, respectively) when joint processing is adopted (Color figure online)

We verify the (mean-field) NE property by testing the deviation of the channel assignment of each of the 19 players. The results shown in Fig. 4 are time-averaged values starting from the slot where the pure strategy can be identified until the end of simulation. It is shown in Fig. 4a that for all players a unilateral deviation produces higher (time-averaged) cost; in other words, the learning algorithm converges to an NE point. This suggests that with the approximated instantaneous reward function in (24) convergence to NE is observed numerically. In addition, we test the change of (time-averaged) capacities under unilateral deviation. As can be seen from Fig. 4b, for most MSs a deviation from the NE strategy reduces their own capacity, which is calculated using (8) and time-averaged. Finally, as depicted in Fig. 4c, there is no significant change on the average capacity when only one player unilaterally deviates from the NE strategy.

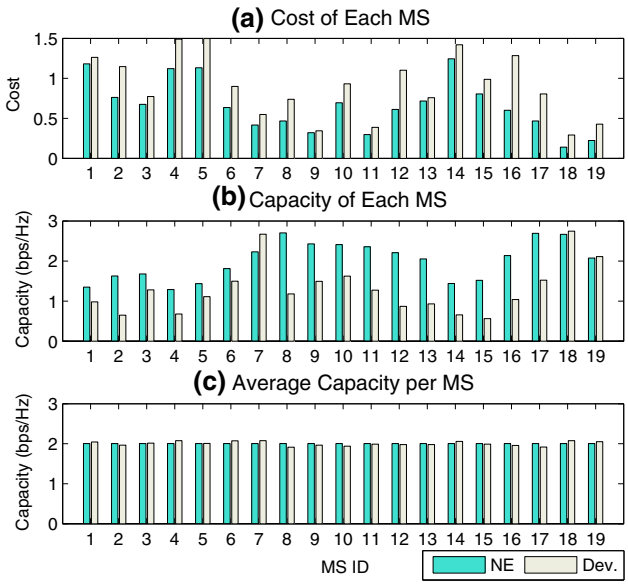


Fig. 4 Cost and capacity for each player for the NE strategy and unilateral deviation from the NE strategy

6.2 Capacity performance for different channel assignment strategies

Here, we compare the capacity performance of the proposed channel assignment strategy with two other methods, namely, the random allocation and centralized selection, which are described as follows:

- In the random allocation scheme, each BS randomly selects a channel for its MS in each frame. No learning algorithm is implemented.
- In the centralized selection scheme, it is assumed that there exists a centralized controller which knows all system information including the channel gains, the channel availability statistics, and the number of BSs. The channel assignment profile is determined by minimizing the total number of mutually interfering links, i.e.,

$$\mathbf{a}_{exh} = \underset{\mathbf{a}\in\mathcal{A}}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{j\in\mathcal{C}_i, j\neq i} \mathbb{1}_{\{a_i=a_j\}}. \tag{32}$$

Figure 5 compares the cumulative distribution function (CDF) of the average cell capacity in each time slot for different channel assignment strategies. As can be seen, the

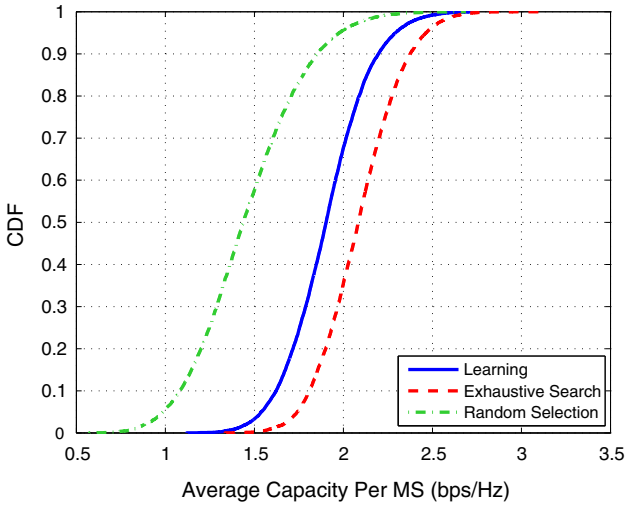


Fig. 5 Comparison of the achievable capacity for three channel assignment strategies when joint processing is adopted

proposed learning algorithm significantly outperforms the random selection approach and performs close to the centralized selection approach. This demonstrates the proposed learning algorithm's ability to allocate mutually interfered players on different channels in its convergence toward the NE point.

6.3 Capacity performance and fairness for different precoding schemes

As mentioned in Sect. 3.2, local precoding is a special case of joint processing. Here, we investigate the impact of different precoding schemes on the performance of the proposed learning algorithm. The average per-MS capacities for different combinations of channel assignment and precoding schemes are summarized in Table 3. For the proposed learning algorithm, it is shown that joint processing yields 10–30 % improvement over local precoding across different channel assignment strategies. The results also suggest that a lower threshold β_{th} will lead to a higher average cell capacity, since when joint processing is adopted nearby cells serve the MS instead of simply mitigating its interference. Besides, we observe an increased capacity gap between the random selection and the centralized selection when joint processing is applied. This is because in joint processing a neighboring BS becomes a serving BS, and when adjacent cells are using the same subchannel the signal for another MS becomes a strong interference source.

In addition to the average per-MS capacity, the fairness among players is examined. Fairness of resource allocation is usually measured by the Jain's fairness index (JFI) [47] which is defined as

Table 3 Capacity per MS (bps/Hz) for different combinations of channel assignment and precoding schemes

Precoding	Learning	Random	Centralized
Local precoding	1.6476	1.5246	1.7006
Joint processing, $\beta_{th} = 0.5$	1.8406	1.6924	1.8993
Joint processing, $\beta_{th} = 0.3$	2.1052	1.8835	2.1811

$$J = \frac{\left(\sum_{i=1}^{N} \bar{R}_i\right)^2}{N \sum_{i=1}^{N} \bar{R}_i^2} \tag{33}$$

where $\bar{R}_i$ is the time-averaged capacity of player i over the whole simulation. The value of JFI falls in $[1/N, 1]$, and a higher JFI value represents better fairness. The JFI of the three channel assignment strategies are summarized in Table 4. As can be seen, the random selection scheme, due to its fully randomized nature, achieves the best fairness in terms of the time-averaged cell capacity while the other two channel assignment strategies are also reasonably fair.

6.4 Performance results for distributed networks with random geometry

The proposed learning algorithm can be implemented in any network with universal frequency reuse. Here, we consider the scenario where the transmission links are randomly placed, which reflects the typical network topology of distributed networks (e.g., cognitive radio and femtocell networks). We generate a topology of 10 links, with the transmitters randomly distributed inside a 1 km by 1 km square area and each receiver located at a distance of 120–150 m away from its transmitter. The transmission power is set to $P_0 = 23$ dBm, with pathloss and shadowing given by the line-of-sight (LOS) urban-micro model [46]. Other simulation parameters follow those in Table 2. A snapshot of the network topology is shown in Fig. 6. Only local precoding is considered in this scenario, since joint processing requires backhaul communications among transmitters, making its implementation difficult in distributed networks.

The evolution of the mixed strategies is depicted in Fig. 7. The convergence toward the pure strategy is clearly observed. In addition, a comparison of different players shows that the convergence behavior is highly related to the interference condition of individual links. For relatively isolated players (e.g., link 9), it takes longer time to converge. In contrast, for players in crowed regions (e.g., links 2, 5, and 6), the convergence is generally faster but with large variation. This can be explained through the proposed reward function. Observe that in the definition in (15), higher interference means higher cost. Thus, the difference between the cost of choosing channels is smaller for isolated links than for links in crowded

Table 4 JFI (33) for different combinations of channel assignment and precoding schemes

Precoding	Learning	Random	Centralized
Local precoding	0.8507	0.9034	0.8530
Joint processing, $\beta_{th} = 0.5$	0.8847	0.9280	0.8809
Joint processing, $\beta_{th} = 0.3$	0.8903	0.9371	0.9034

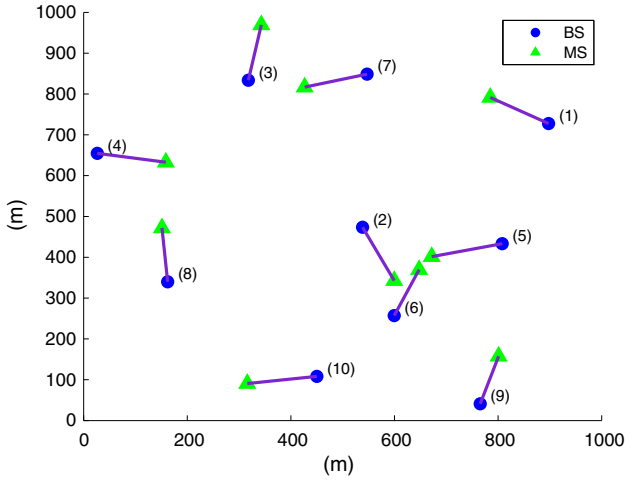


Fig. 6 A snapshot of the nodes' positions and network topology. The link ID is shown in parenthesis next to the link

region. The multiplicative-weights update rule makes a larger probability adjustment in each step in the latter case, resulting in a faster convergence.

Figure 8 shows the convergence behavior of the channel selection by stochastic learning with decreasing learning rates (step sizes), as opposed to constant learning rates in the proposed method. The decreasing learning rates are set to

$$\eta_i(n) = \epsilon_i(n) = \frac{100}{n + 999}, \quad n \geq 1 \tag{34}$$

which start from 0.1 as in the case of constant learning rates. All other parameter settings follow those in Fig. 7. We can see from Figs. 7 and 8 that both learning procedures converge to the same NE point, although the algorithm with decreasing learning rates takes longer to converge.

The performance of the proposed learning algorithm is shown in Fig. 9. Figure 9a compares different channel assignment strategies and shows that the learning algorithm outperforms the random selection. Specifically, for highly interfered users, the proposed algorithm significantly improves the capacity compared to random selection. Comparing centralized selection with the proposed algorithm, we observe their mixed performance across links

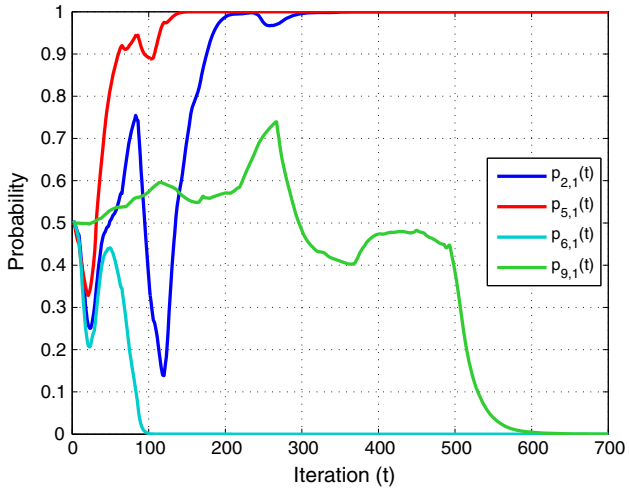


Fig. 7 Evolution of the mixed strategies of four selected players when local precoding is applied to the distributed network

with a comparable average capacity. Note that the proposed learning algorithm finds an NE which coincides with a local maximum of the potential function (i.e., a local minimum of the sum interference), and the centralized selection scheme finds the minimum of the sum number of interfering links. The two objectives are different but aligned with each other, resulting in the comparable performance as numerically demonstrated in Fig. 9a.[2] The test of deviation from the NE property is conducted and the NE property is again verified in Fig. 9b. The increase of cost due to unilateral deviation from NE is significant for highly interfered (crowded) players and slight for isolated players. These observations show that the proposed learning algorithm is effective in networks with random geometry for all kinds of interference conditions.

7 Conclusion

We have studied the problem of distributed channel assignment in multicell network MIMO systems with time-varying channel and unknown number of BSs through a game-theoretic approach. We have proposed a practical joint processing scheme where each MS is jointly served by a set of nearby BSs. We have formulated the channel assignment problem as a noncooperative game where the reward function was properly defined so that the BSs implicitly coordinate their channel assignments. The game was also shown to be an exact potential game. To achieve

[2] Note that Fig. 5 simulates a different topology and shows a slightly different comparison result. However, both Figs. 5 and 9a demonstrate the efficacy of the proposed distributed learning method as compared to a centralized scheme.

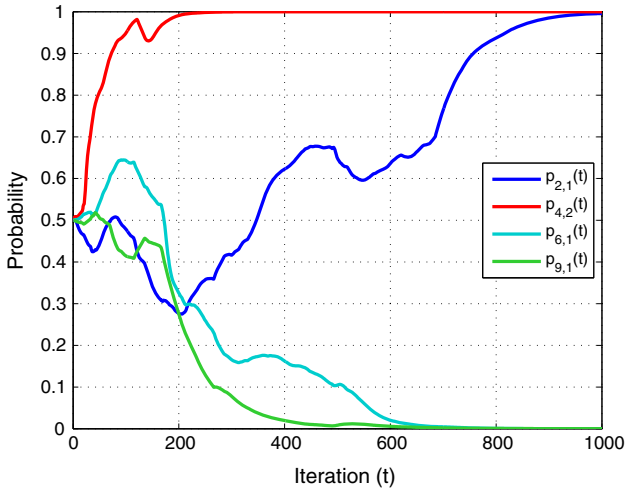


Fig. 8 Evolution of the mixed strategies of four selected players when local precoding is applied to the distributed network, with decreasing learning rates

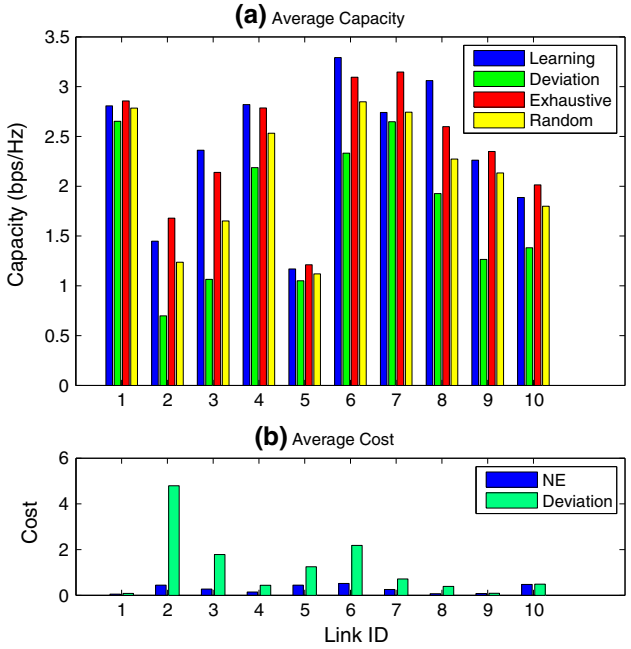


Fig. 9 **a** Achievable capacity for different channel assignment strategies. **b** Cost for each player for the NE strategy and unilateral deviation from the NE strategy

the Nash equilibrium strategy, we have proposed a stochastic learning-based distributed algorithm by which each cell adjusts its channel assignment strategy simultaneously in each iteration without the ordering by any coordinator, according to its action-reward history. The convergence property of the proposed algorithm in achieving an NE point was theoretically proven and numerically verified for different network scenarios. The performance of the proposed algorithm in terms of the achievable capacity and fairness was also examined.

The proposed learning-based channel assignment method has been applied to cellular and distributed networks with multiple antennas. This work may be extended by considering networks consisting of base stations with heterogeneous capabilities in terms of coverage, spectrum, number of antennas, and so on. When heterogeneous base stations are involved in a network, new limiting factors such as different processing capabilities and non-ideal backhaul connections must be considered in the base station cooperation and interactions. While the problem formulation can be more challenging, we believe that the learning-based methodology for distributed channel assignment can be applied in these extended scenarios. Other open challenges include the consideration of mobile stations with high mobility, the impact of imprecise or quantized channel feedback, etc.

Proof of proposition 4

The Proposition is proved by investigating the nondecreasing and upper-bounded properties of Ψ along the trajectory of the ODE in (27). First, we rewrite the ODE in (27) as follows:

$$\frac{dp_{i,s_i}(t)}{dt} = p_{i,s_i}(t) \sum_{s_i' \in \mathcal{A}_i} p_{i,s_i'}(t) \Big[\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \psi_{\mathbf{i}}(\mathbf{e}_{\mathbf{s_i'}}, \mathbf{P}_{-\mathbf{i}}) \Big]. \tag{35}$$

Given that $\Psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) = \partial \Psi(\mathbf{P}) / \partial \mathbf{p}_{\mathbf{i},\mathbf{s_i}}$ is positively correlated with $\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}})$, and let $D_{i,s_i,s_i'} = \psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \psi_i(\mathbf{e}_{\mathbf{s_i'}}, \mathbf{P}_{-\mathbf{i}})$, $E_{i,s_i,s_i'} = \Psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \Psi(\mathbf{e}_{\mathbf{s_i'}}, \mathbf{P}_{-\mathbf{i}})$, we may write

$$D_{i,s_i,s_i'} > 0 \Leftrightarrow E_{i,s_i,s_i'} > 0. \tag{36}$$

By applying (35) and (36), the derivation of $\Psi(\mathbf{P})$ with respect to t is given by

$$\begin{aligned} \frac{d\Psi(\mathbf{P})}{dt} &= \sum_{i \in \mathcal{N}} \sum_{s_i \in \mathcal{A}_i} \frac{\partial \Psi(\mathbf{P})}{\partial p_{i,s_i}} \frac{dp_{i,s_i}}{dt} \\ &= \sum_{i \in \mathcal{N}} \sum_{s_i, s_i' \in \mathcal{A}_i} p_{i,s_i} p_{i,s_i'} \Psi(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) \cdot \mathbf{D}_{\mathbf{i},\mathbf{s_i},\mathbf{s_i'}} \\ &= \frac{1}{2} \sum_{i \in \mathcal{N}} \sum_{\substack{s_i, s_i' \in \mathcal{A}_i \\ s_i < s_i'}} p_{i,s_i} p_{i,s_i'} E_{i,s_i,s_i'} \cdot D_{i,s_i,s_i'} \\ &\geq 0 \end{aligned} \tag{37}$$

where the last inequality holds since given the condition in (36), $D_{i,s_i,s_i'}$ and $E_{i,s_i,s_i'}$ always have the same sign.

Thus, Ψ is nondecreasing along the trajectories of the ODE, and asymptotically all the trajectories will be in the set $\{\mathbf{P} \in \mathcal{P} : \frac{\mathbf{d}\Psi(\mathbf{P})}{\mathbf{dt}} = \mathbf{0}\}$. From (35) and (37), we know

$$\begin{aligned} &\frac{d\Psi(\mathbf{P})}{dt} = 0 \\ &\Rightarrow p_{i,s_i} p_{i,s_i'} \left[\psi_i(\mathbf{e}_{\mathbf{s_i}}, \mathbf{P}_{-\mathbf{i}}) - \psi_{\mathbf{i}}(\mathbf{e}_{\mathbf{s_i'}}, \mathbf{P}_{-\mathbf{i}})\right]^2 = 0, \quad \forall i, s_i, s_i' \\ &\Rightarrow \frac{dp_{i,s_i}}{dt} = 0, \quad \forall i, s_i, s_i' \\ &\Rightarrow \mathbf{P}^* \text{ is a stationary point of the ODE (27).} \end{aligned} \tag{38}$$

According to Proposition 3, when starting from an interior point of the simplex of the mixed strategy space $\mathcal{P}$, the trajectory of the ODE in (35) converges to a stable stationary point, i.e., an NE. Then, by Proposition 2, the linearly interpolated process of the strategy update $p_{i,s_i}(n)$ is bounded within the neighborhood of the trajectory of (35). Thus, we complete the proof [22, Theorem 3.3].

References

1. Zhang, H., Dai, H., & Zhou, Q. (2004). Base station cooperation for multiuser MIMO: Joint transmission and BS selection. In *Proceedings of IEEE CISS '04*.
2. Chang, R. Y., Tao, Z., Zhang, J., & Kuo, C.-C. J. (2009). Multicell OFDMA downlink resource allocation using a graphic framework. *IEEE Transactions on Vehicular Technology*, *58*(7), 3494–3507.
3. Hadisusanto, Y., Thiele, L., & Jungnickel, V. (2008). Distributed base station cooperation via block-diagonalization and dual-decomposition. In *Proceedings of IEEE GLOBECOM '08*, pp. 1–5.
4. Spencer, Q. H., Swindlehurst, A. L., & Haardt, M. (2004). Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Transactions on Signal Processing*, *52*(2), 461–471.
5. Caire, G., Ramprashad, S. A., & Papadopoulos, H. C. (2010). Rethinking network MIMO: Cost of CSIT, performance analysis, and architecture comparisons. In *Proceedings of ITA '10*, pp. 1–10.
6. Zhang, J., Chen, R., Andrews, J. G., Ghosh, A., & Heath, R. W. (2009). Networked MIMO with clustered linear precoding. *IEEE Transactions on Wireless Communications*, *8*(4), 1910–1921.
7. Kaviani, S., Simeone, O., Krzymien, W. A., & Shamai, S. (2011, December). Linear MMSE precoding and equalization for network MIMO with partial cooperation. In *Proceedings of IEEE GLOBECOM '11*, pp. 1–6.
8. de Kerret, P., & Gesbert, D. (2012, April). Sparse precoding in multicell MIMO systems. In *Proceedings of IEEE WCNC '12*, pp. 958–962.
9. de Kerret, P., & Gesbert, D. (2011, August). The multiplexing gain of a two-cell MIMO channel with unequal CSI. In *Proceedings of IEEE ISIT '11*, pp. 558–562.
10. Zakhour, R., Ho, Z., & Gesbert, D. (2009, April). Distributed beamforming coordination in multicell MIMO channels. In *Proceedings of IEEE VTC Spring '09*, pp. 1–5.
11. Zakhour, R., & Gesbert, D. (2010). Distributed multicell-MISO precoding using the layered virtual SINR framework. *IEEE Transactions on Wireless Communications*, *9*(8), 2444–2448.
12. Bjornson, E., Jalden, N., Bengtsson, M., & Ottersten, B. (2011). Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission. *IEEE Transactions on Signal Processing*, *59*(12), 6086–6101.
13. Sundaresan, K., & Rangarajan, S. (2009). Efficient resource management in OFDMA femtocells. In *Proceedings on ACM MobiHoc '09*, pp. 33–42.
14. Lopez-Perez, D., Valcarce, A., de la Roche, G., & Zhang, J. (2009). OFDMA femtocells: A roadmap on interference avoidance. *IEEE Communications Magazine*, *47*(9), 41–48.
15. Hatoum, A., Aitsaadi, N., Langar, R., Boutaba, R., & Pujolle, G. (2011, June). FCRA: Femtocell cluster-based resource allocation scheme for OFDMA networks. In *Proceedings of IEEE ICC '11*, pp. 1–6.
16. Bloem, M., Alpcan, T., & Başar, T. (2007). A stackelberg game for power control and channel allocation in cognitive radio networks. In *Proceedings of ICST VALUETOOLS '07*, p. 4.
17. Husheng, L. (2010). Multiagent Q-learning for aloha-like spectrum access in cognitive radio systems. *EURASIP Journal on Wireless Communications and Networking*, *2010*. http://jwcn.eurasipjournals.com/content/2010/1/876216/.
18. Galindo-Serrano, A., & Giupponi, L. (2011, October) Femtocell systems with self organization capabilities. In *Proceedings of IEEE NetGCoop '11*, pp. 1–7.
19. Nie, N., & Comaniciu, C. (2005, November). Adaptive channel allocation spectrum etiquette for cognitive radio networks. In *Proceedings of IEEE DySPAN '05*, pp. 269–278.
20. Xu, Y., Wang, J., Anpalagan, A., & Yao, Y.-D. (2012). Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution. *IEEE Transactions on Wireless Communications*, *11*(4), 1380–1391.
21. Zhong, W., & Youyun, X. (2010). Game theoretic multimode precoding strategy selection for MIMO multiple access channels. *IEEE Signal Processing Letters*, *17*(6), 563–566.
22. Sastry, P. S., Phansalkar, V. V., & Thathachar, M. A. L. (1994). Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. *IEEE Transactions on Systems, Man and Cybernetics*, *24*(5), 769–777.
23. Tembine, H. (2011). Dynamic robust games in MIMO systems. *IEEE Transactions on Systems, Man and Cybernetics B*, *41*(4), 990–1002.
24. Busoniu, L., Babuska, R., & De Schutter, B. (2008). A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man and Cybernetics C*, *38*(2), 156–172.
25. Cominetti, R., Melo, E., & Sorin, S. (2010). A payoff-based learning procedure and its application to traffic games. *Games and Economic Behavior*, *70*(1), 71–83.
26. Bravo, M. (2011). An adjusted payoff-based procedure for normal form games. arXiv preprint arXiv:1106.5596.
27. Khan, M. A., Tembine, H., & Vasilakos, A. V. (2012). Game dynamics and cost of learning in heterogeneous 4G networks. *IEEE Journal on Selected Areas in Communications*, *30*(1), 198–213.
28. Goldsmith, A. J., & Chua, S.-G. (1997). Variable-rate variable-power MQAM for fading channels. *IEEE Transactions on Communications*, *45*(10), 1218–1230.
29. Sadek, M., Tarighat, A., & Sayed, A. H. (2007). A leakage-based precoding scheme for downlink multi-user MIMO channels. *IEEE Transactions on Wireless Communications*, *6*(5), 1711–1721.
30. Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, *14*, 124–143.

31. Brown, G. W. (1951). Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*, *13*(1), 374–376.
32. Borkar, V. S. (1997). Stochastic approximation with two time scales. *Systems & Control Letters*, *29*(5), 291–294.
33. Leslie, D. S., & Collins, E. J. (2003). Convergent multiple-timescales reinforcement learning algorithms in normal form games. *The Annals of Applied Probability*, *13*(4), 1231–1251.
34. Leslie, D. S., & Collins, E. J. (2005). Individual Q-learning in normal form games. *SIAM Journal on Control and Optimization*, *44*(2), 495–514.
35. Nemirovski, A. S., Juditsky, A., Lan, G., & Shapiro, A. (2009). Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, *19*(4), 1574–1609.
36. Coucheney, P., Gaujal, B., & Mertikopoulos, P. (2014). Penalty-regulated dynamics and robust learning procedures in games. arXiv preprint arXiv:1303.2270.
37. Leslie, D. S., & Collins, E. J. (2006). Generalised weakened fictitious play. *Games and Economic Behavior*, *56*, 285–298.
38. Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games* (Vol. 2). Cambridge: MIT Press.
39. Bournez, O., & Cohen, J. (2013). Learning equilibria in games by stochastic distributed algorithms. In E. Gelenbe & R. Lent (Eds.), *Computer and information sciences III*, (pp. 31–38). London: Springer.
40. Tembine, H. (2012). *Distributed strategic learning for wireless engineers*. Boca Raton: CRC Press.
41. Billingsley, P. (1995). *Probability and measure*. Hoboken: Wiley-Interscience.
42. Kushner, H. J., & Yin, G. G. (2003). *Stochastic approximation and recursive algorithms and applications*. Berlin: Springer.
43. Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. *Séminaire de Probabilités XXXIII*, *1709*, 1–68.
44. Beneveniste, A., Metivier, M., & Priouret, P. (1987). *Adaptive algorithms and stochastic approximations*. Berlin: Springer.
45. Benaïm, M., & Hirsch, M. W. (1999). Stochastic approximation algorithms with constant step size whose average is cooperative. *The Annals of Applied Probability*, *9*(1), 216–241.
46. 3GPP. (2011). *Spatial channel model for multiple input multiple output (MIMO) simulations (release 10)*. 3gpp technical report (tr 25.996) v10.0.0, March 2011.
47. Jain, R., Chiu, D., & Hawe, W. (1984). *A quantitative measure of fairness and discrimination for resource allocation in shared computer systems*. DEC Research Report TR-301.

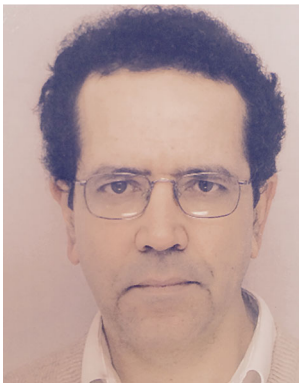
Li-Chuan Tseng received the B.S. and the Ph.D. degree from National Chiao-Tung University, Hsinchu, Taiwan, in 2005 and 2013, respectively, both in electronics engineering. He joined MediaTek Inc., Hsinchu, Taiwan, in 2013. His research interests include cooperative commnication systems, game theoretic resource allocation, and wirless communication standardization.

Feng-Tsun Chien received the B.S. degree from National Tsing Hua University, Hsinchu, Taiwan, in 1995, the M.S. degree from National Taiwan University, Taipei, Taiwan, in 1997, and the Ph.D. degree from the University of Southern California, Los Angeles, CA, USA, in 2004, all in electrical engineering. He joined the Department of Electronics Engineering, National Chiao Tung University, Hsinchu, in July 2005 as an Assistant Professor. His current research interests include wireless communications, statistical signal processing, game theoretic resource allocation, and network information theory. He has been serving as a Technical Program Committee Member in ICC (2009-2015), GLOBECOM (2009-2014), WCNC (2010-2015), and PIMRC (2011-2014) and is currently the treasurer of the IEEE Vehicular Technology Society, Taipei Chapter.

Ronald Y. Chang received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2000, the M.S. degree in Electronics Engineering from National Chiao Tung University, Hsinchu, in 2002, and the Ph.D. degree in Electrical Engineering from the University of Southern California, Los Angeles, CA, USA, in 2008. From 2002 to 2003, he was with the Industrial Technology Research Institute, Hsinchu. In 2008, he was a Research Intern at the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. In 2009, he worked on NASA Small Business Innovation Research projects. Since 2010, he has been with the Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan, where he is currently an Assistant Research Fellow. His research interests include wireless communications and networking. He was an Exemplary Reviewer for the IEEE Communications Letters in 2012, and a recipient of the Best Paper Award from the IEEE Wireless Communications and Networking Conference 2012. He has contributed to various conferences as a Technical Program Committee Member, including the IEEE International Conference on Communications 2012, 2013, and 2015.

Wei-Ho Chung received the B.Sc. and M.Sc. degrees in Electrical Engineering from the National Taiwan University, Taipei, Taiwan, in 2000 and 2002, respectively, and the Ph.D. degree in Electrical Engineering from the University of California, Los Angeles, in 2009. From 2002 to 2005, he was a system engineer at ChungHwa Telecommunications Company, where he worked on data networks. In 2008, he worked on CDMA systems at Qualcomm, Inc., San Diego, CA. His research interests include communications, signal processing, and networks. Dr. Chung received the Taiwan Merit Scholarship from 2005 to 2009 and the Best Paper Award in IEEE WCNC 2012, and has published over 40 journal articles and over 50 conference papers. Since January 2010, Dr. Chung has been an assistant research fellow, and promoted to the rank of associate research fellow in January 2014 in Academia Sinica. He leads the Wireless Communications Lab in the Research Center for Information Technology Innovation, Academia Sinica, Taiwan.

ChingYao Huang received the B.S. degree in physics from National Taiwan University, Taipei, Taiwan, in 1987 and the Master and Ph.D. degrees in Electrical and Computer Engineering from New Jersey Institute of Technology (NJIT), Newark, and Rutgers University, the State University of New Jersey, New Brunswick, in 1991 and 1996, respectively. He joined AT&T, Whippany, NJ, and then Lucent Technologies in 1996 as a Member of Technical Staff. In 2001 and 2002, he was an Adjunct Professor with Rutgers University and NJIT. In 2002, he joined the Department of Electronics Engineering, National Chiao Tung University, HsinChu, Taiwan, where he is currently an Associate Professor and the Director of the Technology Licensing Office and Incubation Center. He has served as Editor for the ACM Wireless Networks and Recent Patents on Electrical Engineering. He has published more than 60 technical memorandums, journal papers, and conference proceeding papers. He is the holder of 16 patents. His research interests include wireless medium access controls for cellular, wireless body area networks, and wireless machine-to-machine communications. Dr. Huang was the Technical Chair for the International Symposium of Medical Information and Communication Technology in 2010. He was the recipient of the Bell Labs Team Award from Lucent Technologies in 2003, the Best Paper Award from the IEEE Vehicular Technology Conference in Fall 2004, and the Outstanding Achievement Award from National Chiao Tung University during 2007-2011.



Abdelwaheb Marzouki received the Ph.D. degree in signal processing from the Université des Sciences et Technologies de Lille, Lille, France. in 1996. He has held both industry and academic positions at Philips Consumer Communications, University of Oulu and Mines-Télécoms. He developed Radar image segmentation algorithms, designed Physical layer modules for UMTS transceiver prototyping, participated in 3gpp standardization activities and edited deliverables for European and national projects. He is the author of several patents in wireless communication and localization systems and published papers dealing with Physical and MAC layer design and Radar image processing. Currently he is assistant professor in Wireless Networks and Multimedia Services Department at Télécom SudParis. His research focus includes resource allocation for MIMO/OFDM wireless systems, interference mitigation techniques for wireless networks and advanced localization techniques.