Distributed Multiple Access with Multiple Transmission Options at The Link Layer

Faeze Heydaryan, Yanru Tang, and Jie Luo

***Abstract*— This paper investigates the problem of distributed medium access control in a wireless multiple access network with an unknown finite number of homogeneous transmitters. An enhanced physical link layer interface is considered where each link layer user can be equipped with multiple transmission options. Assume that each user is backlogged with a saturated message queue. With a generally-modeled channel, a distributed medium access control framework is proposed to adapt the transmission scheme of each user to maximize an arbitrarily chosen symmetric network utility. The framework suggests that the receiver should measure the success probability of a carefully designed virtual packet, and feed such information back to the transmitters. Upon receiving the measured probability, each transmitter should obtain an estimated number of users by comparing the probability with a pre-determined theoretical reference. Transmission schemes of the users are then adapted toward a target that is a function of the estimated number of users. Convergence conditions are characterized for the proposed algorithm to converge to a designed unique equilibrium, which should be close to optimal with respect to the chosen utility. Simulation results are provided to demonstrate the optimality and the convergence properties of the proposed algorithm.**

I. Introduction

Due to increasing dynamics of communication activities, a significant proportion of messages in communication networks are transmitted using distributed protocols where users make their transmission decisions and communication parameter choices individually. Classical network architecture such as the OSI model assumes that each link layer user should be equipped with a single transmission option plus an idling option [1]. At any moment, a link layer user can only choose to idle or to transmit a packet. When communication cannot be fully optimized at the physical layer, which happens often in a distributed wireless network, data link layer must share the responsibility of transmission adaptation. However, the single transmission option setting significantly limited the capability of exploiting advanced wireless tools such as rate and power adaptations at the data link layer.

Recently, a new channel coding theory was proposed in [2][3][4][5] for distributed communication at the physical layer. The coding theory allows each physical layer transmitter to prepare an ensemble of channel codes, and to choose an arbitrary one (according to the link layer decision) to encode its message and to transmit the codeword symbols to the receiver. While code ensembles of the users are assumed to be known, actual coding decisions are not shared among the transmitters or with the receiver. The receiver, on the other hand, should either decode the messages of interest or report collision, depending on whether a pre-determined error probability requirement can be met. Fundamental limit of the system was characterized using a distributed channel capacity region defined in the vector space of the coding decisions of the transmitters. The distributed capacity region was shown in [2][5] to coincide with the classical Shannon capacity region. Error performance bounds in the case of finite codeword length were obtained in [3][4].

The new channel coding theory provided the basic physical layer support for an enhancement to the physical-link layer interface [4][5], which allows each link layer user to be equipped with multiple transmission options. These options correspond to different codes at the physical layer, possibly representing different communication settings such as different transmission power and rate combinations. The interface enhancement enables data link layer protocols to exploit advanced wireless communication adaptations through the navigation of different transmission options. It also enables the modeling of a wide range of general but realistic link layer channels that can be derived from physical layer channel and coding details.

As an early attempt to support the enhanced physical-link layer interface at the link layer, [6][7] investigated the problem of distributed utility optimization in a wireless multiple access network with an unknown finite number of homogeneous users. While a general link layer channel and a general utility function were considered, [6][7] still assumed that each user should be equipped with a single transmission option plus an idling option. Base on a stochastic approximation framework [8][9][10], a distributed medium access control (MAC) algorithm was proposed to adapt the transmission probabilities of the users to a designed unique equilibrium. In the proposed MAC algorithm, the receiver should measure the contention level of the channel using the success probability of a carefully designed virtual packet [6][7]. Once a channel contention estimate is obtained and is fed back to the users (transmitters), each user should then increase/decrease its transmission probability to move channel contention toward a desired theoretical value. It was shown that, at the equilibrium, all users should have the same transmission probability. Transmission probability of the users at the equilibrium can be designed as a closed-form function of the unknown number of users, and can also be set close to optimal with respect

Faeze Heydaryan and Jie Luo are with the Electrical and Computer Engineering Department, Colorado State University, Fort Collins, CO 80523. E-mails: {faezeh66, rockey}@colostate.edu.

Yanru Tang is with Samsung Inc., 9868 Scranton Rd., San Diego, CA 92121. Email: yanru.tang@samsung.com.

This work was supported by the National Science Foundation under Grant CNS-1618960. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

to a chosen network utility. As shown in [6][7], uniqueness of the equilibrium and convergence of the MAC algorithm are guaranteed by the monotonicity properties of two key design functions. On one hand, given the unknown number of users, users can increase/decrease channel contention by decreasing/increasing their transmission probabilities. On the other hand, the expected contention level of the channel should be reverse monotonic in the transmission probability of the users at the equilibrium.

In this paper, we extend the distributed MAC framework of [6][7] to the case when each link layer user can be equipped with multiple transmission options. The key challenge brought by the equipment of multiple transmission options is that, when users adapt their transmission schemes involving a mixture of transmission options, it becomes difficult to determine whether the adaptations should increase or decrease contention level of the channel. Consequently, it becomes mathematically difficult to guarantee the two key monotonicity properties required in [6][7] for the proof of equilibrium uniqueness and convergence of the MAC algorithm. We show in this paper that, such a challenge can be circumvented by a search-assisted design approach. More specifically, the proposed approach uses a manual design part to ensure optimality of the system equilibrium, and uses an automatic design part to guarantee one of the key monotonicity properties required in [6][7]. An alternative proof is adopted to show equilibrium uniqueness and convergence of the MAC algorithm without requiring the other monotonicity property used in [6][7].

The rest of the paper is organized as follows. In Section II, we present a stochastic approximation framework for a class of distributed MAC algorithms with guaranteed convergence to a unique system equilibrium. While the results are more or less standard in the stochastic approximation literature, they characterize the key conditions for convergence and a key approach to simplify the equilibrium analysis. In Section III, based on a general link layer channel model and a utility maximization objective, we present a distributed MAC algorithm that adapts the transmission scheme of each user according to two carefully designed functions. We show that, under a set of assumptions, the proposed MAC algorithm should lead the transmission schemes of all users to a designed system equilibrium. Next, in Section IV, we consider a simple scenario and present a closed-form approach to design the two key functions to satisfy the required assumptions and to place the system equilibrium at a point that is close to optimal with respect to a chosen symmetric network utility[1]. We then extend the design approach to the general scenario in Section V where a search-assisted approach is proposed to replace the closed-form approach to design part of the two key functions. Simulation results are provided in Section VI to demonstrate both the optimality and the convergence properties of the proposed MAC algorithm.

II. A Stochastic Approximation Framework

Consider a distributed multiple access network with a memoryless channel and K homogeneous users (transmitters). Time is slotted. The length of each time slot equals the transmission duration of one packet. We assume that the number of users K should be unknown to the users and also unknown to the receiver. Each user is equipped with M transmission options plus an idling option, and is backlogged with a saturated message queue. We formulate the problem at the data link layer in the sense of constraining users to the provided transmission options. At the beginning of each time slot t, a user should either idle or randomly choose a transmission option to send a message, with corresponding probabilities being specified by an associated probability vector. Transmission decisions of the users are made individually, and they are shared neither among the users nor with the receiver. The M-length probability vector associated to user k, $k = 1, \dots, K$, is denoted by $\boldsymbol{p}_k(t)$ for time slot t. We write $\boldsymbol{p}_k(t) = p_k(t)\boldsymbol{d}_k(t)$, with $0 \le p_k(t) \le 1$ being the probability that user k transmits a packet in time slot t, and with vector $\boldsymbol{d}_k(t)$ specifying the conditional probabilities for user k to choose each of the transmission options should it decide to transmit a packet. Entries of the $\boldsymbol{d}_k(t)$ vector satisfy $0 \le d_{km}(t) \le 1$ for $1 \le m \le M$, and $\sum_{m=1}^{M} d_{km}(t) = 1$. We term $p_k(t)$ the "transmission probability" of user k, and term $\boldsymbol{d}_k(t)$ the "transmission direction" vector of user k.

At the end of each time slot t, based upon available channel feedback, each user k derives a target probability vector $\tilde{\boldsymbol{p}}_k(t)$. User k then updates its transmission probability vector by

$$\begin{aligned}\boldsymbol{p}_k(t+1) &= (1-\alpha(t))\boldsymbol{p}_k(t) + \alpha(t)\tilde{\boldsymbol{p}}_k(t) \\ &= \boldsymbol{p}_k(t) + \alpha(t)(\tilde{\boldsymbol{p}}_k(t) - \boldsymbol{p}_k(t)),\end{aligned} \tag{1}$$

where $\alpha(t) > 0$ is a step size parameter of time slot t. Let $\boldsymbol{P}(t) = [\boldsymbol{p}_1^T(t), \boldsymbol{p}_2^T(t), \dots, \boldsymbol{p}_K^T(t)]^T$ denote an MK-length vector that consists of the transmission probability vectors of all users in time slot t. Let $\tilde{\boldsymbol{P}}(t) = [\tilde{\boldsymbol{p}}_1^T(t), \tilde{\boldsymbol{p}}_2^T(t), \dots, \tilde{\boldsymbol{p}}_K^T(t)]^T$ denote the corresponding target vector. According to (1), $\boldsymbol{P}(t)$ is updated by

$$\boldsymbol{P}(t+1) = \boldsymbol{P}(t) + \alpha(t)(\tilde{\boldsymbol{P}}(t) - \boldsymbol{P}(t)). \tag{2}$$

Probability adaptation given in (2) falls into the stochastic approximation framework [8][9][10], where the target probability vector $\tilde{\boldsymbol{P}}(t)$ is often calculated from noisy estimates of certain system variables, e.g., the channel idling probability.

Define $\hat{\boldsymbol{P}}(t) = [\hat{\boldsymbol{p}}_1^T(t), \hat{\boldsymbol{p}}_2^T(t), \dots, \hat{\boldsymbol{p}}_K^T(t)]^T$ as the "theoretical value" of $\tilde{\boldsymbol{P}}(t)$ under the assumption that there is no measurement noise and no feedback error in time slot t. Let $E_t[\tilde{\boldsymbol{P}}(t)]$ be the conditional expectation of $\tilde{\boldsymbol{P}}(t)$ given system state at the beginning of time slot t. The difference between $E_t[\tilde{\boldsymbol{P}}(t)]$ and $\hat{\boldsymbol{P}}(t)$ is defined as the bias in the target probability vector calculation, denoted by $\boldsymbol{G}(t)$.

$$\boldsymbol{G}(t) = E_t[\tilde{\boldsymbol{P}}(t)] - \hat{\boldsymbol{P}}(t). \tag{3}$$

We assume that, given the communication channel, both $\hat{\boldsymbol{P}}(t) = \hat{\boldsymbol{P}}(\boldsymbol{P}(t))$ and $\boldsymbol{G}(t) = \boldsymbol{G}(\boldsymbol{P}(t))$ should only be functions of $\boldsymbol{P}(t)$, which is the transmission probability vector in time slot t.

The following two conditions are typically required for the convergence of a stochastic approximation algorithm [8][9][10].

[1]A network utility is "symmetric" if it requires that utility values achieved by different individual users should be equal.

Condition 1: (Mean and Bias) There exists a constant $K_m > 0$ and a bounding sequence $0 \leq \beta(t) \leq 1$, such that

$$\|\boldsymbol{G}(\boldsymbol{P}(t))\| \leq K_m \beta(t), \quad (4)$$

where $\|.\|$ denotes the second order norm. We assume that $\beta(t)$ is controllable in the sense that one can design protocols to ensure $\beta(t) \leq \epsilon$ for any chosen $\epsilon > 0$ and for large enough t.

Condition 2: (Lipschitz Continuity) There exists a constant $K_l > 0$, such that

$$\|\hat{\boldsymbol{P}}(\boldsymbol{P}_a) - \hat{\boldsymbol{P}}(\boldsymbol{P}_b)\| \leq K_l \|\boldsymbol{P}_a - \boldsymbol{P}_b\|, \text{ for all } \boldsymbol{P}_a, \boldsymbol{P}_b. \quad (5)$$

According to stochastic approximation theory [8][10], if the above two conditions are satisfied, the step size sequence $\alpha(t)$ and the bounding sequence $\beta(t)$ are small enough, then trajectory of the transmission probability vector $\boldsymbol{P}(t)$ under distributed adaptation given in (2) can be approximated by the following associated ordinary differential equation (ODE) in a sense explained in [8][10],

$$\frac{d\boldsymbol{P}(t)}{dt} = -[\boldsymbol{P}(t) - \hat{\boldsymbol{P}}(t)], \quad (6)$$

where we used t to denote the continuous time variable. Because all entries of $\boldsymbol{P}(t)$ and $\hat{\boldsymbol{P}}(t)$ stay in the range of $[0, 1]$, any equilibrium $\boldsymbol{P}^* = [\boldsymbol{p}_1^{*T}, \ldots, \boldsymbol{p}_K^{*T}]^T$ of the associated ODE must satisfy

$$\boldsymbol{P}^* = \hat{\boldsymbol{P}}(\boldsymbol{P}^*). \quad (7)$$

Suppose that the associated ODE given in (6) has a unique solution at $\boldsymbol{P}^*$, then the following convergence results can be obtained from the standard conclusions in the stochastic approximation literature.

Theorem 1: For distributed transmission probability adaptation given in (2), assume that the associated ODE given in (6) has a unique stable equilibrium at $\boldsymbol{P}^*$. Suppose that $\alpha(t)$ and $\beta(t)$ satisfy the following conditions

$$\sum_{t=0}^{\infty} \alpha(t) = \infty, \sum_{t=0}^{\infty} \alpha(t)^2 < \infty, \sum_{t=0}^{\infty} \alpha(t)\beta(t) < \infty. \quad (8)$$

Under Conditions 1 and 2, $\boldsymbol{P}(t)$ converges to $\boldsymbol{P}^*$ with probability one.

Theorem 1 is implied by [8, Theorem 4.3].

Theorem 2: For distributed transmission probability adaptation given in (2), assume that the associated ODE given in (6) has a unique stable equilibrium at $\boldsymbol{P}^*$. Let Conditions 1 and 2 hold true. Then for any $\epsilon > 0$, there exists a constant $K_w > 0$, such that, for any $0 < \underline{\alpha} < \overline{\alpha} < 1$ satisfying the following constraint

$$\exists T_0 \geq 0, \underline{\alpha} \leq \alpha(t) \leq \overline{\alpha}, \beta(t) \leq \sqrt{\overline{\alpha}}, \forall t \geq T_0, \quad (9)$$

$\boldsymbol{P}(t)$ converges weakly to $\boldsymbol{P}^*$ in the following sense

$$\limsup_{t \to \infty} Pr\left\{\|\boldsymbol{P}(t) - \boldsymbol{P}^*\| \geq \epsilon\right\} < K_w \overline{\alpha}. \quad (10)$$

Theorem 2 can be obtained by following the proof of [10, Theorem 2.3] with minor revisions.

For simplicity, we assumed the same step size sequence $\alpha(t)$ and the same bounding sequence $\beta(t)$ for all users. We also assumed that all users should update their transmission probability vectors synchronously in each time slot. However, according to the literature of stochastic approximation theory [8], convergence results stated in Theorems 1 and 2 should remain valid, even if different users use different step sizes and bounding sequences, so long as the step sizes and bounding sequences of all users satisfy the same constraints given in (8) and (9). Convergence results of Theorems 1 and 2 should also remain valid if users adapt their probability vectors asynchronously, so long as users update their probability vectors frequently enough [8]. Note that information on the asymptotic convergence rate of $\boldsymbol{P}(t) \to \boldsymbol{P}^*$ can be obtained from the eigenvalues of the Hessian matrix $\left.\frac{\partial(\hat{\boldsymbol{P}}(\boldsymbol{P})-\boldsymbol{P})}{\partial \boldsymbol{P}}\right|_{\boldsymbol{P}^*}$ [11]. However, convergence rate discussion is outside the scope of this paper.

Theorems 1 and 2 provided convergence guarantee for a class of distributed MAC algorithms. Within the presented stochastic approximation framework, the key question is how to design a distributed MAC algorithm to satisfy Conditions 1 and 2 and to place the unique equilibrium of the associated ODE at a point that maximizes a chosen utility. Because users are homogeneous, if equilibrium of the system is indeed unique, transmission probability vectors of the users at the equilibrium must be identical. We choose to enforce such a property by guaranteeing that all users should obtain the same target transmission probability vector in each time slot. This is achieved by the following design details [6][7].

We assume that, in each time slot, there is a virtual packet being transmitted through the channel. Virtual packets assumed in different time slots are identical. A virtual packet is an assumed packet whose coding parameters are known to the users and to the receiver, but it is not physically transmitted in the system, i.e., the packet is “virtual”. Without knowing the transmission/idling status of the users, we assume that the receiver can detect whether the reception of a virtual packet should be regarded as successful or not, and therefore can estimate its success probability [4][6][7]. For example, suppose that the link layer channel is a collision channel, and a virtual packet has the same coding parameters of a real packet. Then, virtual packet reception in a time slot should be regarded as successful if and only if no real packet is transmitted. Success probability of the virtual packet in this case equals the idling probability of the collision channel. For another example, if all packets including the virtual packet are encoded using random block codes, given the physical layer channel, reception of the virtual packet corresponds to a detection task that judges whether or not the vector transmission status of all real users should belong to a specific region. Such detection tasks and their performance bounds have been extensively investigated in the distributed channel coding literature [2][3][4][5].

Let $q_v(t)$ denote the success probability of the virtual packet in time slot t. We term $q_v(t)$ the “channel contention measure” because it is designed to serve as a measurement of the contention level of the link-layer multiple access channel. We assume that the receiver should obtain an estimate of $q_v(t)$ and feed it back to all transmitters. Note that, in the collision channel case when $q_v(t)$ equals the channel idling probability, feeding back an estimate of $q_v(t)$ may not be necessary. So long as each user k knows the success probability of its own packet, denoted by $q_k(t)$, idling probability of the channel can be calculated by $(1-p_k(t))q_k(t)$. With a general link layer channel, however, such calculation of $q_v(t)$ at a transmitter is not always possible if an estimate of $q_v(t)$ is not fed back directly by the receiver [6][7]. Upon receiving an estimate of $q_v(t)$, each user calculates its target transmission probability vector as the same function of the $q_v(t)$ estimate. Denote the theoretical target transmission probability vector of a user by $\hat{\boldsymbol{p}}(q_v(t))$. The theoretical target transmission probability vector of all users is given by $\hat{\boldsymbol{P}}(t) = \boldsymbol{1} \otimes \hat{\boldsymbol{p}}(q_v(t))$, where $\boldsymbol{1}$ denotes a K-length vector of all 1's and $\otimes$ represents the Kronecker product. Consequently, according to (6), any equilibrium $\boldsymbol{P}^*$ of the ODE must take the form of $\boldsymbol{P}^* = \boldsymbol{1} \otimes \boldsymbol{p}^*$. Because q_v is a function of the transmission probability vectors of all users, we must have $\boldsymbol{P}^* = \boldsymbol{1} \otimes \boldsymbol{p}^* = \boldsymbol{1} \otimes \hat{\boldsymbol{p}}(\boldsymbol{p}^*)$, where $\hat{\boldsymbol{p}}(\boldsymbol{p}^*)$ denotes the derived target transmission probability vector of a user given that all users have the same transmission probability vector $\boldsymbol{p}^*$.

In a practical system, an estimate of $q_v(t)$ is likely to be corrupted by measurement noise. We assume that, if the transmission probability vectors of all users $\boldsymbol{P}$ is kept at a constant vector, and q_v is measured over an interval of Q time slots, then the measurement should converge to its true value with probability one as Q is taken to infinity. Other than this assumption, measurement noise is not involved in the discussion of the design objectives, i.e., to meet Conditions 1 and 2 and to place the unique system equilibrium at the desired point. Therefore, in the following three sections, we assume that $q_v(t)$ can be measured precisely and be fed back to the users. This leads to $\tilde{\boldsymbol{P}}(t) = \hat{\boldsymbol{P}}(t) = \boldsymbol{1} \otimes \hat{\boldsymbol{p}}(t)$. We will also skip time index t to simplify the notations.

III. Channel Model, Utility, and A Distributed MAC Algorithm

According to the distributed channel coding theory [2][3][4][5], given any combination of transmission status of the users, the receiver should be able to reliably detect the success/failure outcomes of the real and the virtual packets. These outcomes as functions of the transmission status of the users form the complete model of the link layer multiple access channel. While the complete channel model can be overly complicated, we require that the channel should satisfy the following sensitivity assumption.

Assumption 1: (Channel Sensitivity) There exists a finite constant K_c, such that virtual packet reception should fail if the number of parallel real packet transmissions exceeds K_c.

Because it is usually trivial to satisfy this assumption, in the rest of the paper, we will assume that it should hold true.

Given the link layer multiple access channel and the number of users K, channel contention measure $q_v(\boldsymbol{P}, K)$ is a function of the transmission probability vectors of all users $\boldsymbol{P}$. Under Assumption 1, $q_v(\boldsymbol{P}, K)$ equals the summation of a finite number of terms each representing the probability of a particular transmission status combination of the users that can support the successful reception of the virtual packet[2]. Because each of these terms is a polynomial function of $\boldsymbol{P}$, we have the following property.

Theorem 3: With Assumption 1, channel contention measure q_v is Lipschitz continuous in the transmission probability vectors of all users $\boldsymbol{P}$. That is, there exists a finite constant K_{qc}, such that for any number of users K and any transmission probability vectors $\boldsymbol{P}_a$, $\boldsymbol{P}_b$, the following inequality should hold true.

$$|q_v(\boldsymbol{P}_a, K) - q_v(\boldsymbol{P}_b, K)| \leq K_{qc} \|\boldsymbol{P}_a - \boldsymbol{P}_b\|. \quad (11)$$

Proof of the theorem is skipped.

In the rest of the paper, we will simplify the complete link layer channel model into two sets of channel parameter functions, $\{C_{rij}(\boldsymbol{d})\}$ and $\{C_{vj}(\boldsymbol{d})\}$. Assume that all users should have the same transmission direction vector $\boldsymbol{d}$. We define $\{C_{rij}(\boldsymbol{d})\}$ for $1 \leq i \leq M$ and $j \geq 0$ as the "real channel parameter function set". $C_{rij}(\boldsymbol{d})$ is the conditional success probability of a real packet corresponding to the ith transmission option, should the packet be transmitted in parallel with j other real packets. We also define $\{C_{vj}(\boldsymbol{d})\}$ for $j \geq 0$ as the "virtual channel parameter function set". $C_{vj}(\boldsymbol{d})$ is the success probability of the virtual packet should it be transmitted in parallel with j real packets. We assume that $C_{vj}(\boldsymbol{d}) \geq C_{v(j+1)}(\boldsymbol{d})$ should hold for all $j \geq 0$ and for any $\boldsymbol{d}$. That is, with users having the same transmission direction vector $\boldsymbol{d}$, if the number of parallel real packet transmissions increases, the chance of a virtual packet getting through the channel should not increase. Let $\epsilon_v \geq 0$ be a pre-determined constant. We define $J_{\epsilon_v}(\boldsymbol{d})$ as the smallest integer such that $C_{vJ_{\epsilon_v}}(\boldsymbol{d})$ is strictly larger than $C_{v(J_{\epsilon_v}+1)}(\boldsymbol{d}) + \epsilon_v$, i.e.,

$$J_{\epsilon_v}(\boldsymbol{d}) = \arg\min_j C_{vj}(\boldsymbol{d}) > C_{v(j+1)}(\boldsymbol{d}) + \epsilon_v. \quad (12)$$

By definition, $J_{\epsilon_v}(\boldsymbol{d})$ is a function of $\boldsymbol{d}$. Note that the value of ϵ_v needs to be carefully chosen to guarantee the existence of $J_{\epsilon_v}(\boldsymbol{d})$ for all $\boldsymbol{d}$. With Assumption 1, we should have $C_{vj}(\boldsymbol{d}) = 0$ for all $j > K_c$. If the virtual packet is designed properly, we should also have $C_{v0} > 0$, where C_{v0} is not a function of $\boldsymbol{d}$. Therefore, the existence of $J_{\epsilon_v}(\boldsymbol{d})$ is guaranteed if ϵ_v is chosen to satisfy $0 \leq \epsilon_v < \frac{C_{v0}}{K_c}$. Because both $\{C_{rij}(\boldsymbol{d})\}$ and $\{C_{vj}(\boldsymbol{d})\}$ can be derived from the physical layer channel model and the coding parameters of the packets [2][3][4][5], we assume that they should be known at the transmitters and at the receiver. Note that, while $\{C_{vj}(\boldsymbol{d})\}$ depends on the coding detail of the virtual packet, virtual packet is not involved in the calculation of $\{C_{rij}(\boldsymbol{d})\}$.

With the simplified channel model, given the number of users K and under the assumption that all users should have the same transmission probability vector $\boldsymbol{p} = p\boldsymbol{d}$, we write channel contention measure $q_v(\boldsymbol{p}, K)$ as a function of $\boldsymbol{p}$ and K. In this case, $q_v(\boldsymbol{p}, K)$ can be calculated by

$$q_v(\boldsymbol{p}, K) = \sum_{j=0}^{K} \binom{K}{j} p^j (1-p)^{K-j} C_{vj}(\boldsymbol{d}). \quad (13)$$

[2]For example, a particular term can represent the probability that K_0 users idle, K_1 users transmit with the 1st option, K_2 users transmit with the 2nd option, etc, under the constraints that $\sum_{i=0}^{M} K_i = K$ and $\sum_{i=1}^{M} K_i \leq K_c$.

We assume that users intend to maximize a symmetric utility function. Under the assumption that all users should have the same transmission probability vector, the utility function $U(K, \boldsymbol{p}, \{C_{rij}(\boldsymbol{d})\})$ is defined as a function of the number of users K, the common transmission probability vector $\boldsymbol{p} = p\boldsymbol{d}$, and the real channel parameter function set $\{C_{rij}(\boldsymbol{d})\}$. For example, suppose that users intend to maximize the symmetric sum throughput of the network. If the ith transmission option has a communication rate of r_i (bits/time slot), then the utility function should be given by

$$U(K, \boldsymbol{p}, \{C_{rij}(\boldsymbol{d})\}) = K \sum_{i=1}^{M} d_i r_i \sum_{j=0}^{K-1} \binom{K-1}{j} \times p^{j+1}(1-p)^{K-1-j} C_{rij}(\boldsymbol{d}). \quad (14)$$

Next, we present a distributed MAC algorithm that adapts the transmission probability vectors of the users based on the estimated q_v fed back from the receiver and according to two carefully designed functions, both are functions of an estimated number of users $\hat{K}$. The first function is the "theoretical transmission probability vector" function $\boldsymbol{p}^*(\hat{K})$, which denotes the theoretical transmission probability vector of a user if the number of users equals $\hat{K}$. The second function is the "theoretical channel contention measure" function $q_v^*(\hat{K})$, which denotes the theoretical success probability of the virtual packet if the number of users of the system equals $\hat{K}$ and all users have the same transmission probability vector $\boldsymbol{p}^*(\hat{K})$.

Assumption 2: (Estimation Continuity) $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ should be defined for both integer and non-integer $\hat{K}$ values. Their limits as $\hat{K} \to \infty$, denoted by $\boldsymbol{p}^*(\infty) = \lim_{\hat{K}\to\infty} \boldsymbol{p}^*(\hat{K})$ and $q_v^*(\infty) = \lim_{\hat{K}\to\infty} q_v^*(\hat{K})$, should be well defined. For all integer-valued $\hat{K}$, the following equality should be satisfied

$$q_v^*(\hat{K}) = q_v(\boldsymbol{p}^*(\hat{K}), \hat{K}). \quad (15)$$

Assumption 3: (Contention Monotonicity) $q_v^*(\hat{K})$ should be non-increasing in $\hat{K}$. There exists a positive constant $K_{\min}$, such that $q_v^*(\hat{K})$ should be strictly decreasing for $\hat{K} > K_{\min}$, and $\boldsymbol{p}^*(\hat{K})$ should remain a constant vector for $\hat{K} \leq K_{\min}$.

We are now ready to present the distributed MAC algorithm.

Distributed MAC Algorithm:

1) Each user initializes its transmission probability vector.
2) Let $Q > 0$ be a pre-determined integer. Over an interval of Q time slots, the receiver measures the success probability of the virtual packet, denoted by q_v, and feeds q_v back to all users.
3) Upon receiving q_v, each user derives an estimated number of users $\hat{K}$ by solving the following equation.
$$q_v^*(\hat{K}) = q_v, \qquad \text{s.t. } \hat{K} \geq K_{\min}. \quad (16)$$
If a $\hat{K}$ satisfying (16) cannot be found, users set $\hat{K} = K_{\min}$ if $q_v > q_v^*(K_{\min})$, or set $\hat{K} = \infty$ otherwise. Each user then sets the target transmission probability vector at $\hat{\boldsymbol{p}} = \boldsymbol{p}^*(\hat{K})$.
4) Each user, say user k, updates its transmission probability vector by
$$\boldsymbol{p}_k = (1-\alpha)\boldsymbol{p}_k + \alpha\hat{\boldsymbol{p}}, \quad (17)$$
where α is the step size parameter for user k.
5) The process is repeated from Step 2 till transmission probability vectors of all users converge.

To prove convergence of the distributed MAC algorithm, we need two additional assumptions presented below.

Assumption 4: (Target Continuity) Given q_v, let the target transmission probability vector $\hat{\boldsymbol{p}}$ be determined as in Step 3 of the distributed MAC algorithm. $\hat{\boldsymbol{p}}(q_v)$ as a function of q_v should be Lipschitz continuous in q_v. That is, there exists a constant K_{qp}, such that for any q_{v1} and q_{v2}, the following inequality should hold

$$\|\hat{\boldsymbol{p}}(q_{v1}) - \hat{\boldsymbol{p}}(q_{v2})\| \leq K_{qp}|q_{v1} - q_{v2}|. \quad (18)$$

Assumption 5: (Equilibrium Uniqueness) For any number of users $K > K_{\min}$, equation $q_v^*(\hat{K}) = q_v(\boldsymbol{p}^*(\hat{K}), K)$ should have a unique solution at $\hat{K} = K$. For any number of users $K \leq K_{\min}$, equation $q_v^*(\hat{K}) = q_v(\boldsymbol{p}^*(\hat{K}), K)$ should hold for all $\hat{K} \leq K_{\min}$.

Convergence property of the proposed distributed MAC algorithm is stated in the following theorem.

Theorem 4: Consider the K-user multiple access network presented in this section. Under Assumptions 1-5, and with the proposed MAC algorithm, the associated ODE given in (6) has a unique equilibrium at $\boldsymbol{P}^* = \boldsymbol{1} \otimes \boldsymbol{p}^*(K)$. The probability target $\hat{\boldsymbol{p}}(\boldsymbol{P})$ as a function of the transmission probability vectors of all users $\boldsymbol{P}$ satisfies Conditions 1 and 2. Consequently, transmission probability vectors of all users should converge to $\boldsymbol{P}^* = \boldsymbol{1} \otimes \boldsymbol{p}^*(K)$ in the sense specified in Theorems 1 and 2.

Proof of Theorem 4 is given in .

Note that the distributed MAC algorithm guides the adaptation of transmission probability vectors of all users by trying to maintain channel contention measure at an appropriate level. System equilibrium can be designed as a function of the number of users K even though the actual value of K is unknown. While we have not yet provided any optimality argument on how $\boldsymbol{p}^*(\hat{K})$ should be designed to maximize a chosen utility $U(K, \boldsymbol{p}, \{C_{rij}(\boldsymbol{d})\})$, because $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions need to satisfy the required assumptions, it is quite clear that system equilibrium cannot be designed freely.

IV. A Closed-Form Design Approach with Pre-fixed Transmission Direction

In this section, we consider a simple scenario when all users have the same pre-fixed transmission direction vector $\boldsymbol{d}$. We write the theoretical transmission probability vector function $\boldsymbol{p}^*(\hat{K})$ as

$$\boldsymbol{p}^*(\hat{K}) = p^*(\hat{K})\boldsymbol{d}, \quad (19)$$

where $p^*(\hat{K})$ is the "theoretical transmission probability" function that needs to be designed. It is easy to see that the problem becomes equivalent to the case when each user only has a single transmission option, as investigated in [6][7]. We will review the closed-form approach presented in [6][7] to design $p^*(\hat{K})$ and $q_v^*(\hat{K})$ functions to maximize the chosen network utility and to satisfy Assumptions 2-5. Most of the design parameters presented in this section should be functions of $\boldsymbol{d}$. However, for the sake of simple presentation, we will skip $\boldsymbol{d}$ in some of the notations.

With a fixed transmission direction vector $\boldsymbol{d}$, for most of the utility functions of interest, such as the sum throughput function given in (14), an asymptotically optimal solution should keep the expected load of the channel at a constant [11][12]. Therefore, if p_K^* is the optimal transmission probability when the number of users equals K, we should have $\lim_{K\to\infty} Kp_K^* = x^*$ with $x^* > 0$ being obtained by the following asymptotic utility optimization.

$$x^* = \arg\max_x \lim_{K\to\infty} U\left(K, \frac{x}{K}\boldsymbol{d}, \{C_{rj}(\boldsymbol{d})\}\right). \quad (20)$$

Without knowing the actual number of users K, we design $p^*(\hat{K})$ as

$$p^*(\hat{K}) = \min\left\{p_{\max}, \frac{x^*}{\hat{K}+b}\right\}, \quad (21)$$

where $b \geq 1$ is a pre-determined design parameter, and $p_{\max}$ is given by

$$p_{\max} = \min\left\{1, \frac{x^*}{J_{\epsilon_v}(\boldsymbol{d})+b}\right\}, \quad (22)$$

with $J_{\epsilon_v}(\boldsymbol{d})$ being defined in (12). According to Theorem 4, such a design implies that we intend to set system equilibrium at $\boldsymbol{P}^* = \boldsymbol{1} \otimes p^*(K)\boldsymbol{d}$. As shown in [6][7], this equilibrium setting is not only asymptotically optimal as $K \to \infty$, but also often close to optimal for small K values.

The $q_v^*(\hat{K})$ function, on the other hand, can be calculated as $q_v^*(\hat{K}) = q_v(p^*(\hat{K})\boldsymbol{d}, \hat{K})$ for integer-valued $\hat{K}$. For non-integer-valued $\hat{K}$, $q_v^*(\hat{K})$ is designed as follows. Let $N = \lfloor \hat{K} \rfloor$ be the largest integer below $\hat{K}$. Define $q_N(\boldsymbol{p})$ and $q_{N+1}(\boldsymbol{p})$ as

$$q_N(\boldsymbol{p}) = q_v(\boldsymbol{p}, N), \quad q_{N+1}(\boldsymbol{p}) = q_v(\boldsymbol{p}, N+1). \quad (23)$$

$q_v^*(\hat{K})$ is designed as a linear interpretation between $q_N(p^*(\hat{K})\boldsymbol{d})$ and $q_{N+1}(p^*(\hat{K})\boldsymbol{d})$.

$$q_v^*(\hat{K}) = \frac{p^*(\hat{K}) - p^*(N+1)}{p^*(N) - p^*(N+1)} q_N(p^*(\hat{K})\boldsymbol{d}) + \frac{p^*(N) - p^*(\hat{K})}{p^*(N) - p^*(N+1)} q_{N+1}(p^*(\hat{K})\boldsymbol{d}). \quad (24)$$

With $p^*(\hat{K})$ and $q_v^*(\hat{K})$ functions designed in (21) and (24), respectively, Assumption 2 is satisfied. According to the following theorem, Assumption 3 should hold true if design parameter b in (21) is chosen appropriately.

Theorem 5: [6, Theorem 4] Let $x^* > 0$ and $b \geq \max\{1, x^* - \gamma_{\epsilon_v}\}$ with γ_{ϵ_v} being defined as

$$\gamma_{\epsilon_v} = \min_{\hat{N}, \hat{N} \geq J_{\epsilon_v}(\boldsymbol{d}), \hat{N} \geq x^* - b} \frac{\sum_{j=0}^{\hat{N}} j \binom{\hat{N}}{j} \left(\frac{p^*(\hat{N}+1)}{1-p^*(\hat{N}+1)}\right)^j (C_{vj}(\boldsymbol{d}) - C_{v(j+1)}(\boldsymbol{d}))}{\sum_{j=0}^{\hat{N}} \binom{\hat{N}}{j} \left(\frac{p^*(\hat{N}+1)}{1-p^*(\hat{N}+1)}\right)^j (C_{vj}(\boldsymbol{d}) - C_{v(j+1)}(\boldsymbol{d}))}, \quad (25)$$

where $\hat{N}$ only takes integer values. $q_v^*(\hat{K})$ defined in (24) is non-increasing in $\hat{K}$. Furthermore, if $b > \max\{1, x^* - \gamma_{\epsilon_v}\}$ holds with strict inequality, then $q_v^*(\hat{K})$ is strictly decreasing in $\hat{K}$ for $\hat{K} \geq J_{\epsilon_v}(\boldsymbol{d})$.

According to [6, Theorem 4], Assumption 4 should also hold true. Furthermore, because $p^*(\hat{K})$ is non-increasing in $\hat{K}$, given the number of users K, channel contention measure $q_v(p^*(\hat{K})\boldsymbol{d}, K)$ as a function of $\hat{K}$ is non-decreasing in $\hat{K}$. According to [6, Theorem 3], $q_v(p^*(\hat{K})\boldsymbol{d}, K)$ is strictly increasing in $\hat{K}$ for all $K > K_{\min}$ and $\hat{K} \geq \max\{J_{\epsilon_v}(\boldsymbol{d}), x^* - b\}$. Consequently, Assumption 5 should hold true due to the monotonicity properties of $q_v(p^*(\hat{K})\boldsymbol{d}, K)$ and $q_v^*(\hat{K})$.

It is important to note that system design also includes the design of the virtual packet, which affects the virtual channel parameter function set. In the case of a fixed $\boldsymbol{d}$, virtual packet should be chosen to support reasonable sensitivity of channel contention measure to the variation of number of users. As explained in [6, Section 3], a general principle is to choose a virtual packet design such that $J_{\epsilon_v}(\boldsymbol{d})$ and γ_{ϵ_v} are both slightly less than x^* and therefore $b \geq \max\{1, x^* - \gamma_{\epsilon_v}\}$ can take a value close to 1. Also as explained in [6, Section 4], when the receiver does not feedback q_v and each user only knows the success/failure status of its own packets, the distributed MAC algorithm can be revised to use an interpreted channel contention measure and, according to computer simulations, the system can still converge to the same designed system equilibrium.

V. A Search-Assisted Design Approach

In this section, we consider the general scenario when transmission direction vectors of the users are not fixed. To understand the challenges in the design of $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions, we first take a look at a simple example.

Example 1: Consider a time-slotted multiple access network over a multi-packet reception channel. Each user is equipped with two transmission options respectively labeled as the high-rate option and the low-rate option. If all packets are encoded using the low-rate option, then the channel can support the parallel transmissions of no more than 10 packets. We assume that one packet from the high-rate option is equivalent to the combination of 5 low-rate packets. That is, the channel can support the parallel transmissions of n_h high-rate packets plus n_l low-rate packets if and only if $\frac{1}{2}n_h + \frac{1}{10}n_l \leq 1$. The utility function is chosen to be the sum system throughput. Suppose that all users should hold the same transmission probability vector $\boldsymbol{p} = [p_h, p_l]^T$ where p_h and p_l denote the probabilities of a user choosing the high-rate option and the low-rate option, respectively. We obtain the optimum probability vector as $\boldsymbol{p}^* = [p_h^*, p_l^*]^T = \arg\max_{\boldsymbol{p}} U(K, \boldsymbol{p}, \{C_{rij}(\boldsymbol{d})\})$. Figure 1 illustrates p_h^* and p_l^* as functions of the number of users. We can see that, if we write $\boldsymbol{p}^* = p^*\boldsymbol{d}^*$, then $\boldsymbol{d}^*$ is fixed at $\boldsymbol{d}^* = [1, 0]^T$ for $K \leq 2$, and is fixed at $\boldsymbol{d}^* = [0, 1]^T$ for $K \geq 10$. $\boldsymbol{d}^*$ transits from $[1, 0]^T$ to $[0, 1]^T$ in the region of $2 \leq K \leq 10$.

According to the above observation, we assume that the theoretical transmission probability vector function $\boldsymbol{p}^*(\hat{K}) = p^*(\hat{K})\boldsymbol{d}^*(\hat{K})$ should be designed to satisfy the following properties termed the "Head and Tail Condition".

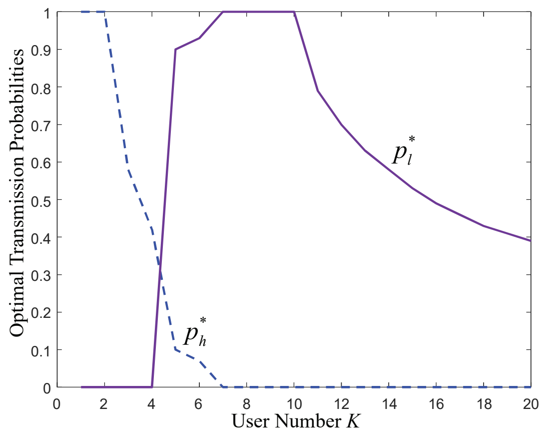

Fig. 1. Optimal transmission probabilities of a K-user multiple access system with each user having two transmission options.

Condition 3: (Head and Tail) Let $\epsilon_v > 0$ be a predetermined constant. Let $J_{\epsilon_v}(\boldsymbol{d})$ be defined in (12). There exist two integer-valued constants $0 < \underline{K} \leq \overline{K}$, such that,

1) $\underline{K} \geq J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$ and $\boldsymbol{d}^*(\hat{K}) = \boldsymbol{d}^*(\underline{K})$ for $\hat{K} \leq \underline{K}$.
2) $\overline{K} > J_{\epsilon_v}(\boldsymbol{d}^*(\overline{K}))$ and $\boldsymbol{d}^*(\hat{K}) = \boldsymbol{d}^*(\overline{K})$ for $\hat{K} \geq \overline{K}$.

The Head and Tail Condition indicates that, when $\hat{K}$ is either small enough or large enough, $\boldsymbol{d}^*(\hat{K})$ should stop changing in $\hat{K}$. Consequently, in the "Head" regime defined as $\hat{K} \leq \underline{K}$, and in the "Tail" regime defined as $\hat{K} \geq \overline{K}$, $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions should be designed using the closed-form approach specified in Section IV.

Now consider the regime of $\underline{K} \leq \hat{K} \leq \overline{K}$. Because we usually have $\boldsymbol{d}^*(\underline{K}) \neq \boldsymbol{d}^*(\overline{K})$, if the designed equilibrium needs to be close to optimal, then theoretical transmission probability vector function $\boldsymbol{p}^*(\hat{K})$ designed for $\underline{K} \leq \hat{K} \leq \overline{K}$ should involve a transition of the transmission direction vector from $\boldsymbol{d}^*(\underline{K})$ to $\boldsymbol{d}^*(\overline{K})$. Unfortunately, due to generality of the system model, when users change their transmission direction vectors, it is difficult to argue whether the outcome should increase/decrease the channel contention measure. Consequently, it becomes difficult to argue for monotonicity properties on channel contention measure functions $q_v(\boldsymbol{p}^*(\hat{K}), K)$ and $q_v^*(\hat{K})$. To overcome such a challenge, we switch to a search-assisted approach whose basic idea is illustrated as follows. We will first choose several integer-valued $\hat{K}$ points, termed "Pinpoints", and assume that $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K}) = q_v(\boldsymbol{p}^*(\hat{K}), \hat{K})$ should be manually determined for the pinpoints. After that, an interpolation approach will be used to connect the pinpoints and to determine $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions for all $\underline{K} \leq \hat{K} \leq \overline{K}$. Key objective of the pinpoints selection and their corresponding design is to make sure that the theoretical transmission probability vector function $\boldsymbol{p}^*(\hat{K})$ is close to optimal in terms of network utility optimization at equilibrium for all $\underline{K} \leq \hat{K} \leq \overline{K}$. Key objective of the interpolation approach, on the other hand, is to make sure that $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions designed to connect the pinpoints should satisfy the monotonicity and continuity requirements presented in Assumptions 2-5.

We require that the following condition should be satisfied by the pinpoints.

Condition 4: (Pinpoints) Let $\hat{K}_i$ for $i = 0, \ldots, L$ be $L+1$ integers such that $\underline{K} = \hat{K}_0 < \hat{K}_1 < \ldots < \hat{K}_L = \overline{K}$. For $i = 0, \ldots, L$ and $0 \leq \lambda < 1$, define

$$\begin{aligned} \hat{K}_{i\lambda} &= (1-\lambda)\hat{K}_{i-1} + \lambda \hat{K}_i, \\ \boldsymbol{d}^*_{i\lambda} &= (1-\lambda)\boldsymbol{d}^*(\hat{K}_{i-1}) + \lambda \boldsymbol{d}^*(\hat{K}_i), \\ q^*_{vi\lambda} &= (1-\lambda)q_v^*(\hat{K}_{i-1}) + \lambda q_v^*(\hat{K}_i). \end{aligned} \quad (26)$$

1) There exists a positive constant ϵ_q to satisfy $q_v^*(\hat{K}_{i-1}) - q_v^*(\hat{K}_i) \geq \epsilon_q$, for all $i = 1, \ldots, L$.
2) There exists a constant $\epsilon_v > 0$, such that for all $i = 1, \ldots, L$ and for all $0 \leq \lambda < 1$, we have $\hat{K}_{i\lambda} > J_{\epsilon_v}(\boldsymbol{d}^*_{i\lambda})$, where $J_{\epsilon_v}(\boldsymbol{d}^*_{i\lambda})$ is defined in (12).
3) There exist $0 < \underline{p} < \overline{p} < 1$ to satisfy $\underline{p} \leq p(\hat{K}_i) \leq \overline{p}$ for all $i = 1, \ldots, L$.
4) Let $N = \lfloor \hat{K} \rfloor$. Define $q_N(\boldsymbol{p})$ and $q_{N+1}(\boldsymbol{p})$ as in (23). Extend the definition of $q_v(\boldsymbol{p}, \hat{K})$ for non-integer-valued $\hat{K}$ as

$$\begin{aligned} q_v(\boldsymbol{p}, \hat{K}) = (N+1-\hat{K})q_N(\boldsymbol{p}) \\ +(\hat{K}-N)q_{N+1}(\boldsymbol{p}), \end{aligned} \quad (27)$$

The following inequality should be satisfied for all $i = 1, \ldots, L$ and for all $0 \leq \lambda < 1$.

$$q_v(\overline{p}\boldsymbol{d}^*_{i\lambda}, \hat{K}_{i\lambda}) \leq q^*_{vi\lambda} \leq q_v(\underline{p}\boldsymbol{d}^*_{i\lambda}, \hat{K}_{i\lambda}). \quad (28)$$

With $\boldsymbol{p}^*(\hat{K})$ being designed for the $L+1$ pinpoints, we propose the following interpolation approach to complete $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions for $\underline{K} \leq \hat{K} \leq \overline{K}$.

Interpolation Approach Assume that $\boldsymbol{p}^*(\hat{K})$ is designed for $\hat{K}_i$, $i = 0, \ldots .L$, with $\underline{K} = \hat{K}_0 < \hat{K}_1 < \ldots < \hat{K}_L = \overline{K}$, to satisfy Condition 4. For $i = 1, \ldots, L$ and for all $0 \leq \lambda < 1$, let $\hat{K}_{i\lambda}$, $\boldsymbol{d}^*_{i\lambda}$ and $q^*_{vi\lambda}$ be defined in (26). Let $q_v(\boldsymbol{p}, \hat{K})$ be defined in (27). We choose $p^*(\hat{K}_{i\lambda})$ to satisfy the following equality.

$$q_v(p^*(\hat{K}_{i\lambda})\boldsymbol{d}^*_{i\lambda}, \hat{K}_{i\lambda}) = q^*_{vi\lambda}. \quad (29)$$

This leads to $\boldsymbol{p}^*(\hat{K}_{i\lambda}) = p^*(\hat{K}_{i\lambda})\boldsymbol{d}^*_{i\lambda}$. Note that the existence of a solution to (29) is guaranteed by Item 4 of Condition 4.

Effectiveness of the Interpolation Approach is stated in the following theorem.

Theorem 6: Assume that $\boldsymbol{p}(\hat{K})$ is designed for a set of $L+1$ pinpoints $\{\hat{K}_i\}$, $i = 0, \ldots, L$, with $\underline{K} = \hat{K}_0 < \hat{K}_1, \ldots, < \hat{K}_L = \overline{K}$, to satisfy Condition 4. After completing the functions using the Interpolation Approach, $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions satisfy Assumptions 2-5 for $\underline{K} \leq \hat{K} \leq \overline{K}$.

The proof of Theorem 6 is given in .

Note that the search-assisted design approach can also be adopted in the simple scenario when either all users have the same pre-fixed $\boldsymbol{d}$ vector or each user has a single transmission option. When there is a noticeable gap between the optimal performance, in terms of network utility maximization at equilibrium, and the performance of the $\boldsymbol{p}^*(\hat{K})$ function designed using the closed-form approach, one can adjust $\boldsymbol{p}^*(\hat{K})$ at carefully selected pinpoints to further improve its optimality. Also note that when users have multiple transmission options, the system should choose a virtual packet design such that channel contention measure is reasonably sensitive to the change of number of users for all transmission option choices.

While we did not provide theoretical guidance on virtual packet design and pinpoint selections for the general scenario, in the next section, we will show that coming up with a reasonably good design should not be a difficult task.

VI. Simulation Results

In this section, we provide computer examples to illustrate both optimality and convergence properties of the proposed MAC algorithm.

Example 2: We will use the system introduced in Example 1 to illustrate the design procedure of the $\boldsymbol{p}^*(\hat{K})$ function when users have multiple transmission options. First, we consider the "Head" and the "Tail" regimes when $\hat{K}$ is either small or large in value. We will add subscript "H" (or "T") to parameters of the "Head" (or the "Tail") regime. Without specifying the values of $\underline{K}$ and $\overline{K}$, we first determine the optimal transmission direction vectors in these two regimes as $\boldsymbol{d}_H = [1,0]^T$ and $\boldsymbol{d}_T = [0,1]^T$. In other words, users should only use the high rate option in the "Head" regime and only use the low rate option in the "Tail" regime. In the "Head" regime, the channel can support the parallel transmissions of no more than 2 high rate packets. The real channel parameter set of the equivalent single option system is given by $\{C_{rj}\}_H$ with $C_{rj} = 1$ for $j \leq 1$ and $C_{rj} = 0$ otherwise. By following the design guideline of Section IV, we get $x_H^* = \arg\max_x (x + x^2)e^{-x} = 1.62$.

We design the virtual packet to be equivalent to the combination of 3 low rate packets. Consequently, the virtual channel parameter set of the equivalent single option system is given by $\{C_{vj}\}_H = \{C_{rj}\}_H$. Choose $\epsilon_v = 0.01$, we get $\gamma_{\epsilon_v H} = J_{\epsilon_v H} = 1$, and $b_H = 1.01$. In the "Tail" regime, on the other hand, the channel can support the parallel transmissions of no more than 10 low rate packets. The real channel parameter set of the equivalent single option system is given by $\{C_{rj}\}_T$ with $C_{rj} = 1$ for $j \leq 9$ and $C_{rj} = 0$ otherwise. This leads to $x_T^* = \arg\max_x \sum_{i=0}^{9} \frac{x^{i+1}}{i!} e^{-x} = 7.30$. Because a virtual packet is equivalent to the combination of 3 low rate packets, virtual channel parameter set of the equivalent single option system in this case is given by $\{C_{vj}\}_T$ with $C_{vj} = 1$ for $j \leq 7$ and $C_{vj} = 0$ otherwise. Therefore, with $\epsilon_v = 0.01$, we have $\gamma_{\epsilon_v T} = J_{\epsilon_v T} = 7$. Luckily, this supports $b_T = 1.01$.

Next, we determine the values of $\underline{K}$ and $\overline{K}$. We first compare two schemes named the "high rate option only" scheme and the "low rate option only" scheme. In the "high rate option only" scheme, we fix $\boldsymbol{d}^*(\hat{K})$ at $[1,0]^T$ for all $\hat{K}$, and set $p^*(\hat{K}) = \min\left\{p_{\max H}, \frac{x_H^*}{\hat{K}+b_H}\right\}$, where $p_{\max H} = \frac{x_H^*}{J_{\epsilon_v H}+b_H}$. In the "low rate option only" scheme, we fix $\boldsymbol{d}^*(\hat{K})$ at $[0,1]^T$ for all $\hat{K}$, and set $p^*(\hat{K}) = \min\left\{p_{\max T}, \frac{x_T^*}{\hat{K}+b_T}\right\}$, where $p_{\max T} = \frac{x_T^*}{J_{\epsilon_v T}+b_T}$. By comparing utility values and theoretical channel contention measures of the two schemes, we choose $\underline{K} = 2$ and $\overline{K} = 11$. Note that we cannot choose a small-valued $\overline{K}$ due to the constraint of $q_v^*(\underline{K}) > q_v^*(\overline{K})$.

Now consider the "Pinpoints Condition" for $\underline{K} \leq \hat{K} \leq \overline{K}$. For transmission direction vectors $\boldsymbol{d}$ satisfying $d_1 > 0$, with a small enough ϵ_v, we generally have $J_{\epsilon_v} = 1$. Therefore, so long as $\boldsymbol{d}^*(\hat{K})$ does not transit too quickly to $[0,1]^T$, the condition of $\hat{K} > J_{\epsilon_v}(\boldsymbol{d}^*(\hat{K}))$ should hold true. Consequently, only two other key conditions need to be satisfied. The first condition is that $q_v^*(\hat{K})$ of the selected pinpoints must be strictly decreasing in $\hat{K}$. The second condition is that $p^*(\hat{K})$ found in the Interpolation Approach should be bounded away from 0 and 1. In addition, from the optimal scheme, we can see that $\boldsymbol{d}^*(\hat{K})$ should transit toward $[0,1]^T$ faster than a linear transition from $\hat{K} = \underline{K}$ to $\hat{K} = \overline{K}$. With these considerations, we choose the following 4 pinpoints. At the edge of the "Head" and the "Tail" regimes, we have $\hat{K}_0 = \underline{K} = 2$ with $\boldsymbol{p}^*(2) = \frac{x_H^*}{\underline{K}+b_H}[1,0]^T$ and $\hat{K}_3 = \overline{K} = 11$ with $\boldsymbol{p}^*(11) = \frac{x_T^*}{\overline{K}+b_T}[0,1]^T$. We also choose other two pinpoints at $\hat{K}_1 = 5$ and $\hat{K}_2 = 6$. We set transmission directions vectors $\boldsymbol{d}^*(5)$ and $\boldsymbol{d}^*(6)$ to be equal to the corresponding optimal transmission direction vectors, i.e., direction vectors extracted from the optimal $\boldsymbol{p}$ vectors that maximize the sum throughput at $K = 5$ and $K = 6$, respectively. Transmission probabilities of these two pinpoints are chosen such that the resulting $q_v^*(\hat{K})$ equals $\frac{\overline{K}-\hat{K}}{\overline{K}-\underline{K}} q_v^*(\underline{K}) + \frac{\hat{K}-\underline{K}}{\overline{K}-\underline{K}} q_v^*(\overline{K})$. Note that, the purpose of designing pinpoints $\hat{K}_1 = 5$ and $\hat{K}_2 = 6$ is to help $\boldsymbol{d}^*(\hat{K})$ to transit appropriately toward $[0,1]^T$. The rest of the $\boldsymbol{p}^*(\hat{K})$ function is completed using the Interpolation Approach for $\underline{K} \leq \hat{K} \leq \overline{K}$. Theoretical channel contention measure $q_v^*(\hat{K})$ of the designed system is illustrated in Figure 2.

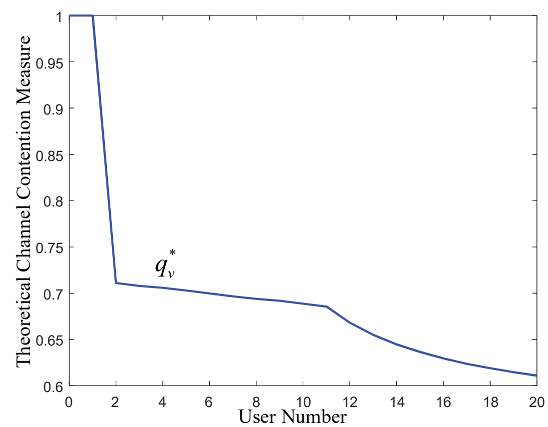

Fig. 2. Theoretical channel contention measure q_v^* as a function of the user number.

In Figure 3, we illustrate the theoretical sum throughput of the network as functions of the number of users K when the transmission probability vectors of all users are set at the following four different vectors: optimal $\boldsymbol{p}(K)$ that maximizes the sum throughput, designed $\boldsymbol{p}^*(K)$, $\boldsymbol{p}^*(K)$ from the high rate option only scheme, and $\boldsymbol{p}^*(K)$ from the low rate option only scheme. Note again that the optimal sum throughput is not necessarily achievable without the knowledge of K. Assume that the high rate only scheme and the low rate only scheme should be reasonably good for the "Head" and the "Tail" regimes, respectively. It can be seen from Figure 3 that, with the help of the designed $\boldsymbol{p}^*(\hat{K})$ and $q_v^*(\hat{K})$ functions, the system can take advantage of the multiple transmission options and maintain a reasonably good performance in term of sum throughput for all user number values.

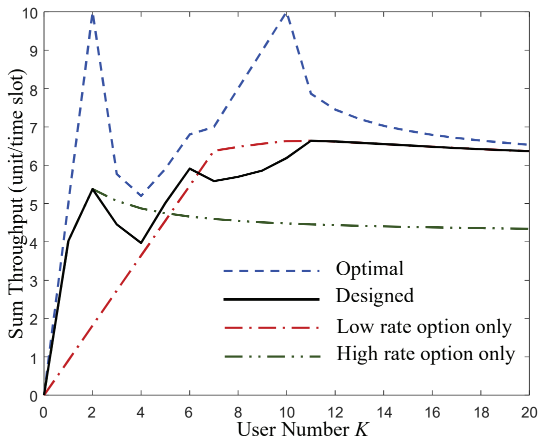

Fig. 3. Sum throughput of the system as functions of the user number under different transmission probability vector settings.

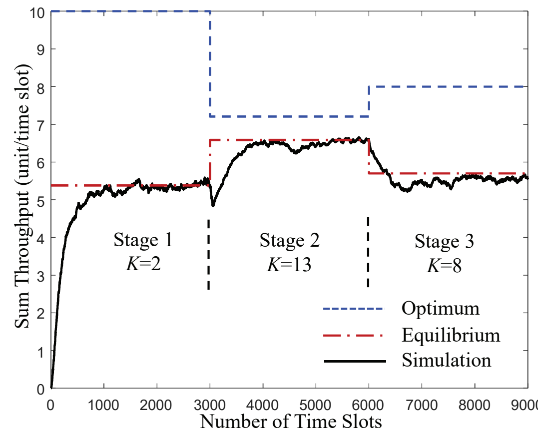

Fig. 4. Convergence in sum throughput of the system. User number changed from 2 to 13 and then to 8 over the three stages.

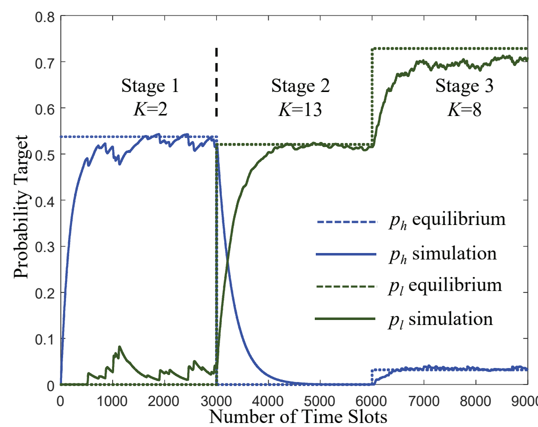

Fig. 5. Entries of the transmission probability vector target and their corresponding theoretical values.

Next, we illustrate the convergence property of the proposed distributed MAC algorithm. Assume that the system has 2 users initially. Transmission probability vectors of all users are initialized at $[0, 0]^T$. In each time slot, according to its own transmission probability vector, each user randomly determines whether a packet should be transmitted or not, and if the answer is positive, which transmission option should be used. The receiver measures q_v using the following exponential moving average approach. q_v is initialized at $q_v = 1$. In each time slot, an indicator variable $I_v \in \{0, 1\}$ is used to represent the success/failure status of the virtual packet reception. q_v is then updated by $q_v = (1 - \frac{1}{300})q_v + \frac{1}{300}I_v$, and is fed back to the users at the end of each time slot. Each user then adapts its transmission probability vector according to the proposed MAC algorithm with a constant step size of $\alpha = 0.05$.

We assume that the system experiences three stages. At Stage one, the system has 2 users. The system enters Stage two at the 3001st time slot, when 11 more users enter into the system with their transmission probability vectors initialized at $[0, 0]^T$. Then at the 6001st time slot, the system enters Stage three when 5 users exit the system. Convergence behavior in sum throughput of the system is illustrated in Figure 4. The corresponding optimal throughput and the theoretical throughput at the designed equilibrium are provided as references. In Figure 5, we also illustrate entries of the target transmission probability vector calculated by the users together with the corresponding theoretical values. Note that the simulated throughput and probability values presented in the figures are measured using the same exponential averaging approach explained above. From Figures 4 and 5, we can see that the proposed MAC algorithm can indeed help users to adapt to the changes of stages and to adjust their transmission probability vectors to the new equilibrium.

According to the Head and Tail Condition, the system degrades to an equivalent single option system when $K \leq \underline{K}$ and $K \geq \overline{K}$. It is generally expected that transmission direction vectors of the "Head" and the "Tail" regimes should be different, i.e., $\boldsymbol{d}(\underline{K}) \neq \boldsymbol{d}(\overline{K})$. In this example, we found one virtual packet design that supports both $b_H = 1.01$ in the "Head" regime and $b_T = 1.01$ in the "Tail" regime. One may think that such a lucky result should be rare. Surprisingly, according to our observations, in most of the problems of interest, even though one may not always be able to get the perfect result of $b_H = b_T \approx 1$, a single virtual packet can often be designed to support close to ideal values on $J_{\epsilon_v H}$, b_H, $J_{\epsilon_v T}$, and b_T. While it is possible to extend the system design and to improve design flexibility by including the transmissions of multiple (different) virtual packets in each time slot, because performance improvement provided by such an extension is often marginal, we choose to skip the corresponding discussions in this paper.

VII. Conclusion

We investigated distributed multiple access networking with an unknown finite number of homogeneous users. An enhanced physical-link layer interface is considered where each link layer user can be equipped with multiple transmission options. With a generally modeled link layer channel, we proposed distributed MAC algorithms to adapt the transmission schemes of the users to maximize a chosen symmetric network utility. Convergence property of the proposed MAC algorithms is proven under quite mild conditions. While there is no theoretical guarantee on the optimality of the proposed MAC algorithms, simulation results suggest that performances of the proposed MAC algorithms are often not too far from optimal.

Appendix

According to Step 3 of the distributed MAC algorithm, users should always have the same target transmission probability vectors. At any equilibrium, we should have transmission probability vectors of all users equal $\boldsymbol{p}^*(\hat{K})$ for some $\hat{K}$, which must satisfy $q_v(\boldsymbol{p}^*(\hat{K}), K) = q_v^*(\hat{K})$. According to Assumption 5, if $K \geq K_{\min}$, we must have $\hat{K} = K$. If $K < K_{\min}$, on the other hand, according to Assumption 3, for all $\hat{K} > K_{\min}$, we have

$$\begin{aligned} q_v^*(\hat{K}) &< q_v^*(K_{\min}) \\ &= q_v(\boldsymbol{p}^*(K_{\min}), K_{\min}) \\ &\leq q_v(\boldsymbol{p}^*(K_{\min}), K). \end{aligned} \quad (30)$$

Consequently, transmission probability vectors of all users at equilibrium must equal $\boldsymbol{p}^*(K_{\min})$, which equals $\boldsymbol{p}^*(K)$ according to Assumption 5. Therefore, the system should always have a unique equilibrium at $\boldsymbol{P}^* = \boldsymbol{1} \otimes \boldsymbol{p}^*(K)$.

Given the number of users K. Target transmission probability vector $\hat{\boldsymbol{p}}$ obtained in Step 3 of the distributed MAC algorithm can be written as a function of the transmission probability vectors of all users $\boldsymbol{P}$ as $\hat{\boldsymbol{p}}(\boldsymbol{P}) = \hat{\boldsymbol{p}}(q_v(\boldsymbol{P}, K))$. Let $\boldsymbol{P}_a$, $\boldsymbol{P}_b$ be two arbitrary transmission probability vectors of all users. According to Assumption 4 and Theorem 3, we have

$$\begin{aligned} &\|\hat{\boldsymbol{p}}(\boldsymbol{P}_a) - \hat{\boldsymbol{p}}(\boldsymbol{P}_b)\| \leq K_{qp}|q_v(\boldsymbol{P}_a, K) - q_v(\boldsymbol{P}_b, K)| \\ &\leq K_{qc}K_{qp}\|\boldsymbol{P}_a - \boldsymbol{P}_b\|. \end{aligned} \quad (31)$$

Therefore, the Lipschitz Continuity Condition 2 is satisfied.

Finally, when the system is noisy, the receiver can choose to measure q_v over an extended number of time slots, or equivalently, to increase the value of Q introduced in Step 2 of the proposed MAC algorithm. If users maintain their transmission probability vectors during the Q time slots, it is often the case that the potential measurement bias in the system can be reduced arbitrarily close to zero. Therefore, the Mean and Bias Condition 1 is also satisfied.

First, it is easy to see that Assumption 2 is satisfied with $q_v^*(\infty)$ being equal to the limiting theoretical channel contention measure of the "Tail" regime, and $\boldsymbol{p}^*(\infty) = p^*(\infty)\boldsymbol{d}^*(\overline{K})$, where $p^*(\infty)$ is the limiting theoretical transmission probability of the "Tail" regime.

Second, in the "Head" regime when $\hat{K} \leq \underline{K}$, because $\underline{K} \geq J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$, $q_v^*(\hat{K})$ is strictly decreasing for $J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K})) \leq \hat{K} \leq \underline{K}$, and $\boldsymbol{p}^*(\hat{K})) = p_{\max}\boldsymbol{d}^*(\underline{K})$ remains a constant vector for $\hat{K} \leq J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$. In other words, we should define $K_{\min} = J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$. Furthermore, $q_v^*(\hat{K})$ is strictly decreasing for $\underline{K} \leq \hat{K} \leq \overline{K}$ by design. Because $\overline{K} > J_{\epsilon_v}(\boldsymbol{d}^*(\overline{K}))$, $q_v^*(\hat{K})$ is also strictly decreasing for $\hat{K} \geq \overline{K}$ in the "Tail" regime. Therefore, Assumption 3 is satisfied.

Third, according to [6, Theorem 4], Assumption 4 should be satisfied in the "Head and Tail" regimes. In other words, target transmission probability vector $\hat{\boldsymbol{p}}(q_v)$ as a function of q_v is Lipschitz continuous in q_v for $q_v \geq q_v^*(\underline{K})$ and $q_v \leq q_v^*(\overline{K})$.

Next, we will prove that the theoretical transmission probability vector function $\boldsymbol{p}^*(\hat{K}) = p^*(\hat{K})\boldsymbol{d}^*(\hat{K})$ is Lipschitz continuous in $\hat{K}$ for $\underline{K} \leq \hat{K} \leq \overline{K}$. Because $\boldsymbol{d}^*(\hat{K})$ is continuous by design, the objective is to show that the search-assisted approach does not lead to any discontinuity of $p^*(\hat{K})$ in $\hat{K}$. For the sake of simple notation, we use $\frac{dp^*(\hat{K})}{d\hat{K}}$ to represent the derivative of $p^*(\hat{K})$ with respect to $\hat{K}$ if $p^*(\hat{K})$ is differentiable. If $p^*(\hat{K})$ is only continuous but not differentiable at $\hat{K}$, then $\frac{dp^*(\hat{K})}{d\hat{K}}$ represents one or an arbitrary subderivative of $p^*(\hat{K})$. If $p^*(\hat{K})$ is not continuous at $\hat{K}$, then $\frac{dp^*(\hat{K})}{d\hat{K}}$ should take the values of $\pm\infty$. Note that the adoption of such a notation does not imply a continuity assumption on $p^*(\hat{K})$. Our objective then becomes to prove that $\frac{dp^*(\hat{K})}{d\hat{K}}$ is bounded for $\underline{K} \leq \hat{K} \leq \overline{K}$.

Let $i \in \{1, \dots, L\}$ and $0 \leq \lambda < 1$ be chosen arbitrarily. Let $\hat{K} = \hat{K}_{i\lambda}$, where $\hat{K}_{i\lambda}$ is defined in (26). To simplify the discussion, we assume that the neighboring two pinpoints satisfy $\hat{K}_{i+1} = \hat{K}_i + 1$, i.e., they take neighboring integer values[3]. Write $\hat{K} = \hat{K}_{i\lambda} = (1-\lambda)\hat{K}_i + \lambda\hat{K}_{i+1}$ as a function of λ, we have $\frac{dp^*(\hat{K})}{d\hat{K}} = \frac{dp^*(\lambda)}{d\lambda}$.

To bound $\frac{dp^*(\lambda)}{d\lambda}$, we consider two different expressions of $q_v^*(\hat{K}) = q_v^*(\lambda)$. The first expression is

$$q_v^*(\lambda) = (1-\lambda)q_v^*(\hat{K}_i) + \lambda q_v^*(\hat{K}_{i+1}). \quad (32)$$

Take derivative with respect to λ, we get $\frac{dq_v^*(\lambda)}{d\lambda} = q_v^*(\hat{K}_{i+1}) - q_v^*(\hat{K}_i)$. Because both $q_v^*(\hat{K}_{i+1})$ and $q_v^*(\hat{K}_i)$ are bounded, there exists a positive constant $\overline{\Delta}_1 > 0$ such that

$$\left|\frac{dq_v^*(\lambda)}{d\lambda}\right| \leq \overline{\Delta}_1. \quad (33)$$

On the other hand, define $p_{i\lambda}^* = p^*(\hat{K}_{i\lambda})$, and consider the second expression of $q_v^*(\hat{K}) = q_v^*(\lambda)$ given below.

$$\begin{aligned} q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*) &= (1-\lambda)q_v(p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*, \hat{K}_i) \\ &+ \lambda q_v(p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*, \hat{K}_{i+1}). \end{aligned} \quad (34)$$

Taking derivative with respect to λ results in

$$\begin{aligned} \frac{dq_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{d\lambda} &= \frac{\partial q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{\partial\lambda} \\ &+ \left[\frac{\partial q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{\partial\boldsymbol{d}_{i\lambda}^*}\right]^T \frac{d\boldsymbol{d}_{i\lambda}^*}{d\lambda} \\ &+ \frac{\partial q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{\partial p_{i\lambda}^*}\frac{dp_{i\lambda}^*}{d\lambda}. \end{aligned} \quad (35)$$

Now we consider the terms on the right hand side of (35) separately.

$$\frac{\partial q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{\partial\lambda} = q_v(p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*, \hat{K}_{i+1}) - q_v(p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*, \hat{K}_i). \quad (36)$$

Because both two terms on the right hand side of (36) are bounded, there exists a constant $\overline{\Delta}_2 > 0$ to satisfy

$$\left|\frac{\partial q_v^*(\lambda, p_{i\lambda}^*\boldsymbol{d}_{i\lambda}^*)}{\partial\lambda}\right| \leq \overline{\Delta}_2. \quad (37)$$

[3]The proof can be easily extended to the case when this assumption does not hold.

According to (13), we can write $q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)$ as

$$q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)=\sum_{j=0}^{\hat{K}_i}\binom{\hat{K}_i}{j}p^{*j}_{i\lambda}(1-p^*_{i\lambda})^{\hat{K}_i-j}C_{vj}(\boldsymbol{d}^*_{i\lambda}). \quad (38)$$

Due to Assumption 1, the right hand side of (38) contains no more than K_c+1 terms. Because $\frac{\partial C_{vj}(\boldsymbol{d}^*_{i\lambda})}{\partial \boldsymbol{d}^*_{i\lambda}}$ is bounded for all j, $\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)}{\partial \boldsymbol{d}^*_{i\lambda}}$ must be bounded. Similarly, $\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_{i+1})}{\partial \boldsymbol{d}^*_{i\lambda}}$ is also bounded. Therefore, from (34), we can see there exists a constant $\overline{\Delta}_3>0$ such that

$$\left|\left[\frac{\partial q^*_v(\lambda,p_{i\lambda}\boldsymbol{d}_{i\lambda})}{\partial \boldsymbol{d}_{i\lambda}}\right]^T\frac{d\boldsymbol{d}_{i\lambda}}{d\lambda}\right|\le\overline{\Delta}_3. \quad (39)$$

From (38), by taking partial derivative with respect to $p^*_{i\lambda}$, we get

$$\begin{aligned}&\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)}{\partial p^*_{i\lambda}}=\\&\sum_{j=0}^{\hat{K}_i}\binom{\hat{K}_i}{j}jp^{*(j-1)}_{i\lambda}(1-p^*_{i\lambda})^{\hat{K}_i-j}C_{vj}(\boldsymbol{d}^*_{i\lambda})\\&-\sum_{j=0}^{\hat{K}_i}\binom{\hat{K}_i}{j}(\hat{K}_i-j)p^{*j}_{i\lambda}(1-p^*_{i\lambda})^{\hat{K}_i-j-1}C_{vj}(\boldsymbol{d}^*_{i\lambda})\\&=\sum_{j=0}^{\hat{K}_i-1}\binom{\hat{K}_i}{j}\hat{K}_ip^{*j}_{i\lambda}(1-p^*_{i\lambda})^{\hat{K}_i-j-1}\\&\times(C_{v(j+1)}(\boldsymbol{d}^*_{i\lambda})-C_{vj}(\boldsymbol{d}^*_{i\lambda})). \quad (40)\end{aligned}$$

Due to Item 2 of the Pinpoints Condition 4, $\hat{K}_i>J_{\epsilon_v}(\boldsymbol{d}^*_{i\lambda})$. Therefore, $C_{vj}(\boldsymbol{d}^*_{i\lambda})-C_{v(j+1)}(\boldsymbol{d}^*_{i\lambda})\ge\epsilon_v$ should hold for at least one $0\le j\le\hat{K}_i-1$. Due to Item 3 of Condition 4, $\underline{p}\le p^*_{i\lambda}\le\overline{p}$. Hence $\left|\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)}{\partial p^*_{i\lambda}}\right|$ is bounded away from zero. The same conclusion applies to $\left|\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_{i+1})}{\partial p^*_{i\lambda}}\right|$. Because both $\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_i)}{\partial p^*_{i\lambda}}$ and $\frac{\partial q_v(p^*_{i\lambda}\boldsymbol{d}^*_{i\lambda},\hat{K}_{i+1})}{\partial p^*_{i\lambda}}$ are negative-valued, from (34), we can see there exists a positive constant $\underline{\Delta}_1>0$ such that

$$\left|\frac{\partial q^*_v(\lambda,p_{i\lambda}\boldsymbol{d}_{i\lambda})}{\partial p_{i\lambda}}\right|\ge\underline{\Delta}_1. \quad (41)$$

Because the two expressions of $q^*_v(\hat{K})$ given in (32) and (34) must equal each other, by combining (33), (35), (37), (39), and (41), we conclude that there exists a positive constant $K_g>0$, such that $\left\|\frac{d\boldsymbol{p}^*(\hat{K})}{d\hat{K}}\right\|\le K_g$. With the extended definition of $\frac{d\boldsymbol{p}^*(\hat{K})}{d\hat{K}}$, as explained at the beginning of the proof, $\left\|\frac{d\boldsymbol{p}^*(\hat{K})}{d\hat{K}}\right\|\le K_g$ means that $\boldsymbol{p}^*(\hat{K})$ is Lipschitz continuous in $\hat{K}$.

According to Item 1 of the Pinpoints Condition 4, for any $\hat{K}_a,\hat{K}_b\in[\underline{K},\overline{K}]$, we have

$$|q^*_v(\hat{K}_a)-q^*_v(\hat{K}_b)|\ge\frac{\epsilon_q}{\overline{K}-\underline{K}}|\hat{K}_a-\hat{K}_b|. \quad (42)$$

This means that, for $\hat{K}$ being obtained according to Step 3 of the distributed MAC algorithm, $\hat{K}(q_v)$ as a function of q_v is Lipschitz continuous in q_v for $q^*_v(\overline{K})\le q_v\le q^*_v(\underline{K})$. Because we just proved that $\boldsymbol{p}^*(\hat{K})$ is Lipschitz continuous in $\hat{K}$, we conclude that the target transmission probability vector $\hat{\boldsymbol{p}}(q_v)$ obtained according to Step 3 of the distributed MAC algorithm is Lipschitz continuous in q_v for $q^*_v(\overline{K})\le q_v\le q^*_v(\underline{K})$. Combined with Lipschitz continuity of $\hat{\boldsymbol{p}}(q_v)$ in the "Head and Tail" regimes, we can see that Assumption 4 is also satisfied.

Fourth, because $q^*_v(\hat{K})$ is strictly decreasing in $\hat{K}$ for $\hat{K}\ge J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$, it is easy to prove that, if $K\le K_{\min}=J_{\epsilon_v}(\boldsymbol{d}^*(\underline{K}))$, then $q^*_v(\hat{K})=q_v(\boldsymbol{p}^*(\hat{K}),K)$ should hold for all $\hat{K}\le K_{\min}$, and for $K_{\min}\le K\le\underline{K}$ and $K\ge\overline{K}$, $q^*_v(\hat{K})=q_v(\boldsymbol{p}^*(\hat{K}),K)$ should have a unique solution at $\hat{K}=K$.

Now consider the case when $\underline{K}\le K\le\overline{K}$. With users setting their transmission probability vectors at $\boldsymbol{p}^*(\hat{K})$, because $\hat{K}>J_{\epsilon_v}(\boldsymbol{d}^*(\hat{K}))$ and $\underline{p}\le p^*(\hat{K})\le\overline{p}$, if $K>\hat{K}$ and $\hat{K}$ is an integer, we must have

$$q_v(\boldsymbol{p}^*(\hat{K}),K)<q_v(\boldsymbol{p}^*(\hat{K}),\hat{K})=q^*_v(\hat{K}). \quad (43)$$

If $K>\hat{K}$ and $\hat{K}$ is not an integer, we have

$$\begin{aligned}&q_v(\boldsymbol{p}^*(\hat{K}),K)<q_v(\boldsymbol{p}^*(\hat{K}),\lfloor\hat{K}\rfloor),\\&q_v(\boldsymbol{p}^*(\hat{K}),K)\le q_v(\boldsymbol{p}^*(\hat{K}),\lfloor\hat{K}\rfloor+1),\end{aligned} \quad (44)$$

which implies that

$$q_v(\boldsymbol{p}^*(\hat{K}),K)<q^*_v(\hat{K}). \quad (45)$$

On the other hand, if $K<\hat{K}$ and $\hat{K}$ is an integer, we must have

$$q_v(\boldsymbol{p}^*(\hat{K}),K)>q_v(\boldsymbol{p}^*(\hat{K}),\hat{K})=q^*_v(\hat{K}). \quad (46)$$

If $K<\hat{K}$ and $\hat{K}$ is not an integer, we have

$$\begin{aligned}&q_v(\boldsymbol{p}^*(\hat{K}),K)>q_v(\boldsymbol{p}^*(\hat{K}),\lfloor\hat{K}\rfloor+1),\\&q_v(\boldsymbol{p}^*(\hat{K}),K)\ge q_v(\boldsymbol{p}^*(\hat{K}),\lfloor\hat{K}\rfloor),\end{aligned} \quad (47)$$

which also implies that

$$q_v(\boldsymbol{p}^*(\hat{K}),K)>q^*_v(\hat{K}). \quad (48)$$

Consequently, $q^*_v(\hat{K})=q_v(\boldsymbol{p}^*(\hat{K}),K)$ must have a unique solution at $\hat{K}=K$. Therefore, Assumption 5 should hold true.

References

[1] D. Bertsekas and R. Gallager, *Data Network,* 2nd Ed., Prentice Hall, 1992.

[2] J. Luo and A. Ephremides, A New Approach to Random Access: Reliable Communication and Reliable Collision Detection, IEEE Trans. Inform. Theory, Vol. 58, pp. 379-423, Feb. 2012.

[3] Z. Wang and J. Luo, Error Performance of Channel Coding in Random Access Communication, IEEE Trans. Inform. Theory, Vol. 58, pp. 3961-3974, Jun. 2012.

[4] J. Luo, A Generalized Channel Coding Theory for Distributed Communication, IEEE Trans. Commun., Vol. 63, pp. 1043-1056, Apr. 2015.

[5] Y. Tang, F. Heydaryan, and J. Luo, Distributed Capacity of A Multiple Access Channel, IEEE WCNC, Marakech, Morocco, Apr. 2019.

[6] Y. Tang, F. Heydaryan, and J. Luo, Distributed Multiple Access with A General Link Layer Channel, Ad Hoc Networks, Vol. 85, pp. 120-130, Mar. 2019.

[7] Y. Tang, F. Heydaryan, and J. Luo, Distributed MAC over a General Multi-packet Reception Channel, WiOPT, Avignon, France, Jun. 2019.

[8] H. Kushner and G. Yin, Stochastic Approximation Algorithms and Applications, New York: Springer-Verlag, 1997.
[9] S. Karlin and H. Taylor, A First Course in Stochastic Processes, 2nd Ed. Academic, San Diego, CA, 1975.
[10] V. Borkar and S. Meyn. The O.D.E Method for Convergence of Stochastic Approximation and Reinforcement Learning, SIAM J. Contrl. and Opt., Vol. 38, pp. 447-469, Jan. 2000.
[11] B. Hajek, Stochastic Approximation Methods for Decentralized Control of Multiaccess Communications, IEEE Trans. Inform. Theory, Vol. IT-31, pp. 176-184, Mar. 1985.
[12] S. Ghez, S. Verdú, and S. Schwartz, Stability Properties of Slotted ALOHA with Multipacket Reception Capability, IEEE Trans. Auto. Contrl., Vol. 33, pp. 640-649, Jul. 1988.