Proportional Fair MU-MIMO in 802.11 WLANs

V. Valls and D. J. Leith

Abstract—We consider the proportional fair rate allocation in an 802.11 WLAN that supports multi-user MIMO (MU-MIMO) transmission by one or more stations. We characterise, for the first time, the proportional fair allocation of MU-MIMO spatial streams and station transmission opportunities. While a number of features carry over from the case without MU-MIMO, in general neither flows nor stations need to be allocated equal airtime when MU-MIMO is available.

Index Terms-MU-MIMO, 802.11ac, proportional fairness.

I. INTRODUCTION

THE next generation of 802.11 WLANs are expected L to support MU-MIMO transmission, whereby parallel transmissions can be simultaneously made to multiple stations. This significantly extends the MIMO support introduced by the 802.11n standard and is, for example, included as part of the current draft 802.11ac standard that aims to support wireless data rates in excess of 1 Gbps. MU-MIMO offers much greater flexibility in scheduling MIMO transmissions, but immediately raises the question of how best to allocate MIMO spatial streams amongst network flows so as to balance fairness and performance. In this paper we consider the proportional fair allocation in an 802.11 WLAN that supports MU-MIMO transmission by one or more stations. While proportional fairness [1] in WLANs has been the subject of considerable interest in the literature, it has only recently been put on a rigorous basis in [2] and consideration of MU-MIMO is new.

The main contribution of the paper is to rigorously characterise, for the first time, the proportional fair allocation of spatial streams and station transmission opportunities in WLANs where one or more stations support MU-MIMO. We demonstrate that this allocation can be found in a distributed manner, without the need for message passing. We show that a number of features carry over from the case without MU-MIMO, specifically that the rate region boundary is characterised by the station total airtimes summing to unity and that stations carring the same number of flows then stations are assigned equal total airtime. Importantly, however, we find that MU-MIMO generally leads to a qualitatively different allocation of airtime compared to the situation without MU-MIMO. Namely, in general flows are not allocated equal airtime. This is because it is the station total airtime that corresponds to the shared network resource being consumed and so to the "cost" of transmissions. When MU-MIMO

transmission is available, flow transmissions occur in parallel and so multiple flows can share the same station airtime.

II. RELATED WORK

In [3] the authors propose a novel MAC design with opportunistic MU-MIMO scheduling based on channel sounding feedback, where packets are selected depending on their transmission duration and type of traffic. In [4] is also proposed a novel MAC design for MU-MIMO that focusses on issues such as MAC ACKing of MU-MIMO transmissions. Packets are scheduled for transmission using a weighted queuing mechanism that considers both packets acknowledgements and type of traffic. However, in both [3] [4] fairness and allocation of MU-MIMO transmission patterns amongst flows is not considered. The work in [5] focusses on packet aggregation in an IEEE 802.11ac AP, and considers a fixed MU-MIMO schedule where one flow is allocated per spatial stream. Regarding utility fairness in WLANs, in [2] is presented the first rigorous analysis of proportional fairness in 802.11 WLANs where transmissions are to a single destination.

III. NETWORK MODEL

A. Preliminaries

We take as our starting point the network model in [2]. Consider an 802.11 WLAN with *n* stations, where each station *i* attempts to transmit at each MAC slot with probability τ_i . We will assume that stations are configured with $CW_{min} = CW_{max}$ (which is supported in 802.11 starting with 802.11e/WME in 2005, 802.11-2007 and subsequent standards), so that the attempt probability is independent of the success or failure of the last transmission. Moreover, it is also assumed that there are no hidden terminals, so all nodes in the network can sense any ongoing transmission. Because of this, a collision can only happen if two or more stations transmit in the same slot. We also assume that noise-induced losses are negligible, although this assumption could be relaxed.

The probability that a transmission by station *i* is successful is the probability that only station *i* transmits and is given by $P_{succ,i} = \tau_i \prod_{k=1,k\neq i}^n (1-\tau_k)$. The probability that a MAC slot is idle is given by the probability that none of the stations transmit, $P_{idle} = \prod_{k=1}^n (1-\tau_k)$. Finally, the probability that a transmission by station *i* collides is $P_{coll,i} = 1 - P_{succ,i} - P_{idle}$. The throughput of station *i* is then given by

$$S_i(\boldsymbol{\tau}) = \frac{P_{succ,i}D_i}{\sigma P_{idle} + T_s(1 - P_{idle})},\tag{1}$$

where σ is the duration of an idle slot, T_s the duration of a busy slot (either successful or collision) and D_i is the size in bits of the frame payload of station *i*. Throughput model (1) is standard, see e.g. [2], [6], [7] and references therein.

Manuscript received December 20, 2013. The associate editor coordinating the review of this letter and approving it for publication was J.-C. Chen.

This material is based upon works supported by the Science Foundation Ireland under Grant No. 11/PI/1177. The authors are with the Hamilton Institute, NUI Maynooth (e-mail:

victor.valls@nuim.ie).

Digital Object Identifier 10.1109/WCL.2014.020314.130884

Let $x_i = \tau_i/(1 - \tau_i)$, thus $x_i \in [0, \infty)$ as $\tau_i \in [0, 1]$. With this change of variable we have that $P_{idle} = 1/\prod_{k=1}^{n} (1+x_k)$, $P_{succ,i} = x_i P_{idle}$, and

$$S_i(\mathbf{x}) = \frac{x_i}{X(\mathbf{x})} \frac{D_i}{T_s},\tag{2}$$

where $X(\mathbf{x}) = a + \prod_{k=1}^{n} (1+x_k) - 1$ with $a = \sigma/T_s$ and $\mathbf{x} = [x_1, \ldots, x_n]$. Notice that $x_i/X(\mathbf{x})$ is the successful airtime for station *i*, and D_i/T_s the rate. Hence, the total airtime (T_i) of station *i* is given by the airtime spent on successful transmissions and collisions

$$T_i = \frac{x_i}{X(\mathbf{x})} \left(1 + \frac{P_{coll,i}}{1 - P_{coll,i}} \right).$$
(3)

B. Extension to MU-MIMO

The throughput model (1) can be extended as follows to encompass MU-MIMO, where stations can transmit multiple spatial streams simultaneously. Let F_i be the set of flows carried by station i, and $F = \bigcup_{i=1}^n F_i$ the set of flows in the WLAN. We let vector \mathbf{v}_{ik} describe the k^{th} MU-MIMO transmission pattern on station *i*, where \mathbf{v}_{ik} has $|F_i|$ elements, and element v_{ikf} defines the number of spatial streams allocated to flow f in this pattern. We collect the set of K_i possible transmission patterns for station *i* together to form matrix \mathbf{V}_i , where the k^{th} row of \mathbf{V}_i describes the k^{th} pattern, $k = 1, ..., K_i$. See for example Figure 1. The set of allowable transmission patterns will be determined by the network characteristics, i.e. number of antennas of the stations, channel conditions and protocol constraints. For example, the draft 802.11ac restricts the use of MU-MIMO to the AP and allows at most 8 spatial streams with at most 4 streams for one client station. However, to keep our analysis as general as possible we will not make any assumptions about the structure of matrix V.



Fig. 1. Example of MU-MIMO transmission matrix V_i where each row represents a possible MU-MIMO transmission pattern for station *i*. For example, row 1 defines a pattern where two spatial streams are allocated to flow f_1 , and one spatial stream each to flows f_3 and f_4 .

Next, let π_{ik} denote the fraction of transmission opportunities that pattern k is selected by station i, with $\sum_{k=1}^{K_i} \pi_{ik} = 1$. We collect the π_{ik} for station i together in vector π_i . We can then express the throughput of flow f on station i as

$$s(f) = \frac{x_i}{X(\mathbf{x})} \frac{D_f}{T_s} \sum_{k=1}^{K_i} \pi_{ik} v_{ikf}, \qquad (4)$$

where $\sum_{k=1}^{K_i} \pi_{ik} v_{ikf}$ is the average number of spatial streams used by flow f in station i, $D_f \sum_{k=1}^{K_i} \pi_{ik} v_{ikf}$ is the average number of bits sent for a flow f in a successful transmission and D_f is the number of bits transmitted by flow f on a single spatial stream in a successful transmission. Note that since spatial streams are transmitted in parallel, a MU-MIMO transmission occupies the same amount of airtime as a single spatial stream and so the total airtime T_i used by station *i* is still given by (3).

In (4) the number of bits transmitted (D_f) by flow $f \in F_i$, i = 1, ..., n does not depend on the selected transmission pattern. However, due to varying channel conditions (such as inter-user interference) it is likely that the number of bits transmitted by flow f varies with the transmission pattern used. Our model can be easily extended to include this. Let matrix $\mathbf{D}_i \in \mathbb{R}_+^{K_i \times |F_i|}$, i = 1, ..., n denote the average number of bits of flow $f \in F_i$ in a transmission pattern $k = 1, ..., K_i$. Then, the throughput expression of a flow f is given by

$$s(f) = \frac{x_i}{X(\mathbf{x})} \frac{1}{T_s} \sum_{k=1}^{K_i} \pi_{ik} v_{ikf} d_{ikf}, \qquad (5)$$

Nevertheless, as this generalisation is straightforward, we use (4) rather than (5) for the rest of the analysis to streamline notation.

IV. PROPORTIONAL FAIR RATE ALLOCATION

A. Log-convexity

It can be readily verified that the flow throughput (4) is nonconvex in x, and also in τ . Fortunately, however, we have the following:

Lemma 1 (Log-convexity):

$$-\tilde{x}_i - \log\left(\sum_{k=1}^{K_i} \pi_{ik} v_{ikf}\right) - \log\frac{D_f}{T_s} + \log X(e^{\tilde{\mathbf{x}}}) \quad (6)$$

is convex in $\tilde{\mathbf{x}}$ and $\boldsymbol{\pi}$, where $\tilde{\mathbf{x}} = [\tilde{x}_1, \dots, \tilde{x}_n]^T$, $\tilde{x}_i = \log x_i$.

Proof: Observe that the first term is linear in \tilde{x} (and so convex), the second term is convex in π due to the convexity of the negative log function when composed with a linear map [8]. The last term is convex in \tilde{x} by Lemma 1 of [2].

B. Utility-fair optimisation

The proportional fair rate allocation is the solution to the utility-fair optimisation problem P:

$$\max_{\tilde{\mathbf{x}},\tilde{\mathbf{s}},\boldsymbol{\pi}} \quad \sum_{f \in F} \tilde{s}(f) \tag{7}$$

s.t.
$$\tilde{s}(f) \leq \log\left(\frac{e^{\tilde{x}_i}}{X(e^{\tilde{\mathbf{x}}})}\frac{D_f}{T_s}\sum_{k=1}^{K_i}\pi_{ik}v_{ikf}\right), \ f \in F_i$$
 (8)

$$\sum_{k=1}^{i} \pi_{ik} = 1, \qquad \pi_{ik} \ge 0, \ i = 1, \dots, n$$
(9)

where $\tilde{s}(f) = \log s(f), \ \tilde{x}_i = \log x_i$.

It follows from Lemma 1 that constraint (8) is convex. Since the objective and remaining constraints are linear in the transformed variables, the optimisation problem is convex and so a proportional fair allocation exists. The proportional fair rate allocation is almost completely characterised as follows:

Theorem 1 (Proportional Fairness): The MU-MIMO proportional fair rate allocation is characterised by: (i) the airtime allocated to station i is $T_i = \frac{|F_i|}{|F|}$ where $|F_i|$ is the number

of flows carried by station *i* and |F| the total number of flows in the WLAN, (ii) the station total airtimes sum to unity $\sum_{i=1}^{n} T_i = 1$, (iii) the allocation of MU-MIMO transmission patterns on station *j* satisfies

$$\sum_{f \in F_j} \lambda_f \frac{v_{jlf}}{\sum_{k=1}^{K_j} \pi_{jk} v_{jkf}} = \nu_j - \theta_{jl}, \ l = 1, \dots, K_j \quad (10)$$

where ν_j , θ_{jl} j = 1, ..., n, $l = 1, ..., K_j$ are non-negative multipliers.

Proof: Optimisation problem P is convex and satisfies the Slater condition, hence strong duality holds. The Lagrangian is

$$\mathcal{L}(\tilde{\mathbf{x}}, \tilde{\mathbf{s}}, \boldsymbol{\pi}, \boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\Theta}) = \sum_{f \in F} \tilde{s}(f) + \sum_{i=1}^{n} \sum_{f \in F_i} \lambda_f \left(\log \frac{e^{\tilde{x}_i} \sum_{k=1}^{K_i} \pi_{ik} v_{ikf}}{X(e^{\tilde{\mathbf{x}}})} \frac{D_f}{T_s} - \tilde{s}(f) \right) + \sum_{i=1}^{n} \nu_i \left(1 - \sum_{k=1}^{K_i} \pi_{ik} \right) + \sum_{i=1}^{n} \sum_{k=1}^{K_i} \theta_{ik} \pi_{ik}$$

where multipliers $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_{|F|}]^T$, $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n]^T$ and $\boldsymbol{\nu} = [\nu_1, \dots, \nu_n]^T$ with $\boldsymbol{\theta}_i = [\theta_{i1}, \dots, \theta_{i|K_i|}]^T$. The main KKT conditions are:

$$\lambda_f = 1,\tag{11}$$

$$\sum_{f \in F_j} \lambda_f - \sum_{i=1}^n \left(\frac{x_j}{X(\mathbf{x})} \prod_{k=1, k \neq j}^n (1+x_k) \sum_{f \in F_i} \lambda_f \right) = 0, \quad (12)$$

$$\sum_{F \in F_j} \lambda_f \frac{v_{jlf}}{\sum_{k=1}^{K_j} \pi_{jk} v_{jkf}} = \nu_j - \theta_{jl}.$$
(13)

Claim (i): From the second KKT condition (12), substituting $\lambda_f = 1$ and rearranging terms we obtain

$$\frac{|F_j|}{|F|} = \frac{x_j}{X(\mathbf{x})} \left(1 + \frac{P_{coll,j}}{1 - P_{coll,j}} \right) =: T_j \tag{14}$$

provided $|F| \neq 0$. Claim (ii) that $\sum_{i=1}^{n} T_i = 1$ follows immediately from (14). Claim (iii) follows from the third KKT condition (13).

Note that property (ii), that station airtimes sum to unity, in Theorem 1 extends to MU-MIMO WLANs the result in [9] that this airtime constraint characterises the WLAN rate region boundary.

C. Determining station transmission attempt probability

Determining the station transmission attempt rates x requires meeting the constraint that the sum of airtimes sums to unity, and so requires knowledge of all station airtimes in the WLAN. However, as discussed in [2], decentralised approximations can be found based on local observations of channel idle time.

D. Determining ν_j , θ_{jl}

The proportional fair rate allocation depends on multipliers ν_j , θ_{jl} . These depend on the distribution of flows amongst the wireless stations, and on the available MU-MIMO transmission patterns at each station and so cannot be stated in closed-form. However, they can be readily determined using standard sub-gradient methods. Namely, by iterating on update

 $\nu_j(t+1) = \nu_j(t) + \alpha(1 - \sum_{k=1}^{K_i} \pi_{ik}), \theta_{jl}(t+1) = \theta_{jl}(t) + \alpha \pi_{ik},$ where $\alpha > 0$ is a sufficiently small step-size parameter. Since these updates make use only of information which is locally available at station *j* they can be implemented in a fully decentralised manner (with no need for message passing).

E. Determining the MU-MIMO transmission pattern

The proportional fair transmission pattern conditions (10) can be expressed in matrix form as

$$\mathbf{V}_{j}(\mathbf{V}_{j}^{T}\boldsymbol{\pi}_{j})^{-*} = \nu_{j}\mathbf{1} - \boldsymbol{\theta}_{j}, \ j = 1, \dots, n$$
(15)

where $\mathbf{x}^{-*} := [\frac{1}{x_1}, \dots, \frac{1}{x_n}]^T$ for vector $\mathbf{x} = [x_1, \dots, x_n]^T$ and 1 denotes the all ones column vector.

When \mathbf{V}_j has full column rank $|F_j|$ (this is commonly satisfied e.g. when the set of possible transmit patterns admits the option to transmit each flow separately in which case \mathbf{V}_j contains the $|F_j| \times |F_j|$ identity matrix), then we can write $\mathbf{V}_j := \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$ where \mathbf{X} is full rank and the rows of \mathbf{Y} are linear combinations of the rows of \mathbf{X} . This partitioning can always be achieved simply by ordering the rows of \mathbf{V}_j appropriately. Condition (15) becomes

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} (\mathbf{V}_j^T \boldsymbol{\pi}_j)^{-*} = \nu_j \mathbf{1} - \boldsymbol{\theta}_j.$$
(16)

Premultiplying both sides by $\begin{bmatrix} \mathbf{X}^{-1} & \mathbf{0} \end{bmatrix}$ and re-arranging,

$$\mathbf{V}_{j}^{T}\boldsymbol{\pi}_{j} = \left(\begin{bmatrix} \mathbf{X}^{-1} & \mathbf{0} \end{bmatrix} (\nu_{j}\mathbf{1} - \boldsymbol{\theta}_{j}) \right)^{-*}.$$
 (17)

Given ν_j and θ_j , vectors π_j satisfying (17) can be found using gaussian elimination. When \mathbf{V}_j is non-singular, then the solution to (17) is unique. However, in general more than one such vector will exist and any such vector is proportional fair. The RHS of (17) depends only on the multipliers associated with wireless station j, which as already noted can be determined using information available locally at station j. Hence, (17) can be solved to find the proportional fair MU-MIMO transmission patterm for each station j in a fully decentralised manner.

F. Finite load

Optimisation problem P can be extended to include flow finite offered loads by adding an additional constraint $\tilde{s}_f \leq \bar{s}$ for each flow f, where \bar{s} is the maximum offered load for flow f. Since these constraints are linear, the optimisation problem remains convex and the foregoing analysis can be directly extended.

V. EXAMPLES

A. Unequal airtimes with MU-MIMO

Consider a WLAN downlink with a MU-MIMO equipped AP that carries 4 flows (f_1, f_2, f_3, f_4) transmitted to four client stations. The offered load is unconstrained. The matrix of available MU-MIMO transmission patterns at the AP is

$$\mathbf{V} = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 2 & 0 & 0 & 1 \\ 2 & 2 & 2 & 0 \\ 1 & 0 & 4 & 2 \end{bmatrix}.$$
 (18)

Since the AP is the only transmitter, the optimal $P_{coll,AP} = 0$, $\tau_{AP} = 1$, and AP total airtime $T_{AP} = 1$. Solving optimisation problem (P), the proportional fair allocation of MU-MIMO transmission patterns is $\pi = \left[\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3}\right]^T$.

The proportional fair rate allocation often corresponds to an equal airtime allocation. The appropriate definition to use for flow airtime is not clear when MU-MIMO is used. One option is the total airtime that would be needed by flow f in order to obtain the same throughput when using a single spatial stream, which is given by $T_i \sum_{k=1}^{K_i} \pi_{ik} v_{ikf}$ and is proportional to the average number of spatial streams allocated to the flow. In the present example, this is 1 for flow 1 and 2 for flows 2, 3 and 4. Another option is the fraction of station i airtime $\sum_{k=1}^{K_i} \pi_{ik} v_{ikf} / \sum_{f \in F_i} \sum_{k=1}^{K_i} \pi_{ik} v_{ikf}$ used by flow f spatial streams in this example is 0.1429 for flow 1 and 0.2857 for flows 2, 3 and 4. A third option is the fraction of station is the station transmission opportunities at which a flow transmits, and in this example we have that each flow is scheduled in 2/3 of the transmissions.

Importantly, observe that none of these flow airtimes are equal at the proportional fair allocation. This is because it is the *station* total airtime that corresponds to the shared network resource being consumed and so to the "cost" of transmissions. Indeed this is reflected in Theorem 1. When MU-MIMO transmission is available, flow transmissions occur in parallel and so multiple flows can share the same station airtime. For a given station airtime, the proportional fair allocation of spatial streams maximises the sum of log flow rates, and this need not correspond to allocating the same number of spatial streams or the same fraction of transmission opportunities to flows.

B. IEEE 802.11ac with Rayleigh fading

We extend the previous example to make the proportional fair allocation of transmission patterns depend on the network characteristics. Consider the WLAN set up of Example V-A with an IEEE 802.11ac AP, channel bandwidth of 20 MHz and guard interval of 800 ns. For simplicity we assume that all spatial streams use BPSK 1/2 modulation and coding scheme, and that the transmission power is equally divided amongst the spatial streams in a transmission pattern. We further assume for simplicity that the fading is independent for each antenna in the WLAN and that the AP has full knowledge of the channel. We consider two types of schedulers, proportional fair and uniform, i.e. transmission patterns are allocated the same fraction of transmission opportunities. Regarding the channel we use Rayleigh fading.

Notice that differently from Example V-A, now in optimisation problem P we have to use the rate as in (5) in order to take into account the network characteristics. That is, matrix **D** in (5) depends on the SNR because it contains the number of bits that can be transmitted for each flow in each transmission pattern. See in Figure 2 how the sum of log flow rates depends on the SNR, scheduler and fading. Next, observe in Figure 3 how the proportional fair allocation of the transmission patterns, and so the flow's airtimes, changes depending on the SNR and channel characteristics. Moreover, notice that the proportional fair allocation converges to the solution of Example V-A when the SNR is large enough.



Fig. 2. Sum of log flow rates in Example V-B for the proportional fair and uniform schedulers with and without Rayleigh fading.



Fig. 3. Fraction of time that transmission patterns are selected in Example V-B for the proportional fair scheduler with and without Rayleigh fading. Transmission patterns correspond to the rows of matrix \mathbf{V} . Transmission pattern 2 is never selected.

VI. CONCLUSIONS

We consider the proportional fair rate allocation in an 802.11 WLAN that supports MU-MIMO transmission by one or more stations. We characterise, for the first time, the proportional fair allocation of MU-MIMO spatial streams and station transmission opportunities. While a number of features carry over from the case without MU-MIMO, in general neither flows nor stations need to be allocated equal airtime when MU-MIMO is available.

REFERENCES

- K. F. P., M. A. K., and T. D. K. H., "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998.
- [2] A. Checco and D. Leith, "Proportional fairness in 802.11 wireless LANs," IEEE Commun. Lett., vol. 15, no. 8, pp. 807–809, Aug. 2011.
- [3] L. Cai, H. Shan, W. Zhuang, X. Shen, J. Mark, and Z. Wang, "A distributed multi-user MIMO MAC protocol for wireless local area networks," in *Proc. 2008 GLOBECOM*, pp. 1–5.
- [4] M. Gong, E. Perahia, R. Stacey, R. Want, and S. Mao, "A CSMA/CA MAC protocol for multi-user MIMO wireless LANs," in *Proc. 2010 GLOBECOM*, pp. 1–6.
- [5] B. Bellalta, J. Barcelo, D. Staehle, A. Vinel, and M. Oliver, "On the performance of packet aggregation in IEEE 802.11ac MU-MIMO WLANs," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1588–1591, 2012.
- [6] D. Leith, Q. Cao, and V. Subramanian, "Max-Min fairness in 802.11 mesh networks," *IEEE/ACM Trans. Networking*, vol. 20, no. 3, pp. 756–769, 2012.
- [7] D. Malone, K. Duffy, and D. Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Trans. Networking*, vol. 15, no. 1, pp. 159–172, 2007.
- [8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [9] V. G. Subramanian and D. J. Leith, "Convexity Conditions for 802.11 WLANs," *IEEE Trans. Inf. Theory*, in press, 2013.