Matrix Exponential Learning: Distributed Optimization in MIMO systems

Panayotis Mertikopoulos, E. Veronica Belmega, and Aris L. Moustakas

***Abstract*—We analyze the problem of finding the optimal signal covariance matrix for multiple-input multiple-output (MIMO) multiple access channels by using an approach based on "exponential learning", a novel optimization method which applies more generally to (quasi-)convex problems defined over sets of positive-definite matrices (with or without trace constraints). If the channels are static, the system users converge to a power allocation profile which attains the sum capacity of the channel exponentially fast (in practice, within a few iterations); otherwise, if the channels fluctuate stochastically over time (following e.g. a stationary ergodic process), users converge to a power profile which attains their *ergodic* sum capacity instead.**

An important feature of the algorithm is that its speed can be controlled by tuning the users' learning rate; correspondingly, the algorithm converges within a few iterations even when the number of users and/or antennas per user in the system is large.

***Index Terms*—Distributed optimization; exponential learning; multiple access channel; MIMO; sum rate.**

I. Introduction

FOLLOWING the seminal prediction that the use of multiple-input multiple-output (MIMO) technologies in signal transmission and reception can lead to substantial performance gains [1], [2], MIMO has become an integral component of numerous state-of-the-art wireless protocols (4G, HSPA+, 802.11n WiFi and WiMAX to name but a few). As a result, considerable impetus has been afforded to developing distributed algorithms that would allow the users of a MIMO system to attain their performance limits at a network level.

On that account, since the actual theoretical limits of MIMO models still elude us even in basic network models (such as the interference channel), it is useful to start instead with the mutual information for Gaussian input and noise, and to optimize the input covariance matrix of each transmitter in the presence of interference from other network users. In this way, one obtains a nonlinear (and possibly non-convex) optimization problem defined over a set of positive-definite matrices, representing the users' power allocation policies (i.e. the spread of their symbol distributions over their antennas). However, given the non-explicit nature of the problem's constraints, standard gradient descent or interior point methods do not apply, so these problems are usually solved by means of the classical waterfilling algorithm [3], properly adapted to multi-user environments [4].

P. Mertikopoulos is with the Centre National de la Recherche Scientifique (CNRS) and the Laboratoire d'Informatique de Grenoble; E.V. Belmega is with ETIS/ENSEA – Université de Cergy-Pontoise – CNRS; A. L. Moustakas is with the Physics Department of the University of Athens, Greece. The authors gratefully acknowledge financial support from the Pôle de Recherche en Mathématiques, Sciences et Technologies de l'Information et de la Communication under grant no. C-UJF-LACODS MSTIC 2012.

The most recent and general incarnations of waterfilling [5] are fully distributed and thus apply to large, unregulated networks where users cannot be assumed to adhere to central control. Unfortunately however, the convergence of these algorithms typically depends on the channel satisfying certain "mild-interference" conditions [6] which, quite often, fail to hold: in fact, in the simple case of a single receiver and several transmitters who communicate over non-overlapping channels (the parallel multiple access channel (PMAC) model), it was shown that these conditions *always* fail [7]. Furthermore, if the channels are not static but evolve over time following a stationary ergodic process (e.g. due to fading), then the problem becomes significantly more difficult (see e.g. [8] for a survey or [9] for some recent results in the asymptotic regime).

Instead of taking a waterfilling approach, we present here an optimization method for problems defined over sets of positive-definite matrices which works by tracking the gradient of the objective function in an *unconstrained* space, and then maps the resulting orbits back to the original (constrained) state space via exponentiating (an approach similar to the one used in [10] for learning a positive semidifinite matrix online). In this manner, if the function to be minimized is convex, we show that this method of "exponential learning" converges exponentially fast to a global minimum (Theorem 1).

Obviously, this method applies to a wide array of MIMO problems, but for concreteness, we will focus on the MIMO multiple access channel (MAC) with transmit power constraints. In this setting, we show that exponential learning converges to the system's maximum achievable sum rate, and this convergence is *independent* of whether the channels are static or ergodic (Theorem 2). More importantly, the speed of this convergence can be controlled by a learning rate parameter, allowing the system to equilibrate within a few iterations even for very large numbers of users and/or antennas per user.

II. System Model

We begin by considering a vector Gaussian multiple access channel with a finite set $\mathcal{K} = \{1, \ldots, K\}$ of K wireless users, each of whom transmits simultaneously over m_k antennas to an n-antenna receiver who decodes individual messages by treating the signals of other users as interference. More specifically, this corresponds to the familiar baseband signal model:

$$\mathbf{y} = \textstyle\sum_k \mathbf{H}_k \mathbf{x}_k + \mathbf{z}, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^n$ is the aggregate message reaching the receiver, $\mathbf{x}_k \in \mathbb{C}^{m_k}$ is the individual message transmitted by user $k \in \mathcal{K}$, $\mathbf{H}_k \in \mathbb{C}^{n \times m_k}$ is the corresponding $n \times m_k$ channel matrix and

$\mathbf{z} \in \mathbb{C}^n$ is the channel noise (assumed zero-mean Gaussian, and without loss of generality, with identity covariance matrix).

In this context, the total transmit power of user k is simply $\mathbb{E}[\|\mathbf{x}_k\|^2] = \operatorname{tr}(\mathbf{P}_k)$, where the expectation is taken over the codebook of user k (assumed Gaussian), and $\mathbf{P}_k$ denotes the covariance matrix $\mathbf{P}_k = \operatorname{cov}(\mathbf{x}_k) = \mathbb{E}[\mathbf{x}_k\mathbf{x}_k^\dagger]$. As is then customary, the performance metric that we will be using is the system's achievable sum rate, i.e. the maximum information transmission rate for a given set of covariance matrices.

This objective naturally depends on the variability of the channel matrices $\mathbf{H}$ over time, so we will consider two different (and diametrically opposed) scenarios:

a) Static Channels: the channel matrices $\mathbf{H}_k$, $k \in \mathcal{K}$ are drawn randomly at the outset of the transmission but remain fixed for its duration. In this case, the system's sum rate is [2]:

$$\Phi(\mathbf{P}) = \log\det\left(\mathbf{I} + \textstyle\sum_k \mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^\dagger\right), \tag{2}$$

where $\mathbf{P}$ denotes the collective profile $\mathbf{P} = (\mathbf{P}_1, \dots, \mathbf{P}_K)$.

b) Fast-fading Channels: in the presence of fast fading, the channel matrices $\mathbf{H}_k$ are stationary ergodic processes with a characteristic time-scale much faster than the typical transmission block (for simplicity, we will also assume that they are temporally uncorrelated).[1] Under these assumptions, we have the following expression for the users' achievable sum rate [12]:

$$\overline{\Phi}(\mathbf{P}) = \mathbb{E}[\Phi(\mathbf{P})], \tag{3}$$

where the expectation is now taken over the variables $\mathbf{H}_k$.

In both the static and the ergodic case, higher transmit powers lead to higher sum rates (individually at least), so users can be assumed to saturate their power constraints.In this way, we obtain the optimization problem:

$$\begin{aligned} &\text{minimize} && F(\mathbf{P}),\\ &\text{subject to} && \mathbf{P}_k \succcurlyeq 0,\ \operatorname{tr}(\mathbf{P}_k) = P_k \quad (k = 1, \dots, K), \end{aligned} \tag{MP}$$

where $F = -\Phi$ or $F = -\overline{\Phi}$ depending on the channel model, and the power levels P_k are non-negative real numbers.

As is well-known, Φ and $\overline{\Phi}$ are both concave, so F is convex; furthermore, if we denote each user's state space by $\mathcal{X}_k = \{\mathbf{P}_k \in \mathbb{C}^{m_k\times m_k} : \mathbf{P}_k \succcurlyeq 0,\ \operatorname{tr}(\mathbf{P}_k) = P_k\}$, then it is easy to see that the problem's state space $\mathcal{X} \equiv \prod_k \mathcal{X}_k$ of (MP) is also convex (viewed as a subset of the complex space $\mathbb{C}^Q$, $Q = \sum_k m_k^2$), making (MP) itself convex. Our goal will thus be to present a general solution method for problems of the type (MP), which when restricted to the objectives (2)-(3), will allow the users of the channel to attain their sum capacity.[2]

It is important to remark here that any solution of (MP) with respect to the sum rate objectives (2)-(3) is also *individually optimal* in the sense that users cannot improve their individual rates by unilaterally changing their power matrices $\mathbf{P}_k$. Indeed, under the single user decoding (SUD) scheme in which the receiver treats the signal of all other users as interference, the achievable rate of user $k \in \mathcal{K}$ for static channels is just:

$$u_k(\mathbf{P}) = \log\det\left(\mathbf{I} + \mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^\dagger\mathbf{W}_k^{-1}\right), \tag{4}$$

where $\mathbf{W}_k = \mathbf{I} + \sum_{\ell\neq k}\mathbf{H}_\ell\mathbf{P}_\ell\mathbf{H}_\ell^\dagger$ is the interference-plus-noise matrix for user k; on the other hand, for ergodically fluctuating channels, we have the expression:

$$\overline{u}_k(\mathbf{P}) = \mathbb{E}[u_k(\mathbf{P})]. \tag{5}$$

The sum rate functions (2)–(3) obviously do not correspond to the sums of (4)–(5) over all transmitters $k \in \mathcal{K}$, but, as was shown in [13], any solution $\mathbf{Q}$ of (MP) with the static objective (2) (resp. the ergodic objective (3)) will satisfy:

$$u_k(\mathbf{Q}) \geq u_k(\mathbf{Q}'_k; \mathbf{Q}_{-k}) \quad (\text{resp. } \overline{u}_k(\mathbf{Q}) \geq \overline{u}_k(\mathbf{Q}'_k; \mathbf{Q}_{-k})), \tag{6}$$

for all $k \in \mathcal{K}$, i.e. *it will be at Nash equilibrium*.[3] In other words, users are aligned with their global objective in the MIMO multiple access channel, so solving (MP) is both globally and individually optimal: *even selfish users have nothing to gain by unilaterally deviating from the global optimum of (MP).*

III. Exponential Learning

The main challenge in solving (MP) is that the positivity constraints $\mathbf{P}_k \succcurlyeq 0$ cannot be expressed in functional form, so (Lagrangian) descent or interior point methods do not readily apply; instead, the static MIMO problem (2) is usually solved by the well-known method of waterfilling [4], [5], [9]. Our aim here will be to overcome this difficulty and present an interior point method which *does* apply to the problem, in both the static and ergodic incarnation of Eqs. (2) and (3) respectively.

A. Exponential learning in parallel multiple access channels

One special case of (MP) which *can* be solved by a variant gradient descent method is the so-called "parallel MAC" where the channels do not overlap and the matrices $\mathbf{P}_k$ are diagonal: $\mathbf{P}_k = \operatorname{diag}(p_{k,1}, \dots, p_{k,m_k})$, with $p_{k\alpha} \geq 0$ and $\sum_{\beta=1}^{m_k} p_{k\beta} = P_k$. In this setting, the system's configuration space $\mathcal{X}$ is a product of simplices, and if we let $v_{k\alpha} = -\frac{\partial F}{\partial p_{k\alpha}}$ denote (minus) the gradient of F, then all interior orbits of the dynamics

$$\dot{p}_{k\alpha} = p_{k\alpha}\left(v_{k\alpha} - P_k^{-1}\textstyle\sum_{\beta=1}^{m_k} p_{k\beta}v_{k\beta}\right), \quad \alpha = 1, \dots, m_k, \tag{7}$$

converge to the minimum of F exponentially fast [11].[4]

This dynamical system is known in game theory and biology as the *replicator equation*, and it is one of the most well-studied models for the evolution of biological populations under natural selection [15], [16]. Regrettably however, this approach cannot be extended to the MIMO case because there is no obvious way to treat (7) as a matrix equation. On the other hand, the replicator dynamics (7) can also be derived from the "exponential learning" scheme [17]:

$$\dot{y}_{k\alpha} = v_{k\alpha} \tag{8a}$$

$$p_{k\alpha} = P_k\frac{\exp(y_{k\alpha})}{\sum_{\beta=1}^{m_k}\exp(y_{k\beta})}, \tag{8b}$$

[1] In fact, the time-uncorrelated case can be seen as a worst-case scenario: temporally correlated models (such as Jakes fading) typically yield much faster convergence times because the channels evolve at a smoother pace [11].

[2] We should stress here that the trace constraint will not be important in our analysis; in fact, it can be easily replaced by any number of (convex) functional constraints of the form $G(\mathbf{P}) = 0$ without affecting the validity of our results.

[3] Recall here that the shorthand $(\mathbf{Q}'_k; \mathbf{Q}_{-k})$ stands for $(\mathbf{Q}_1, \dots, \mathbf{Q}'_k, \dots, \mathbf{Q}_K)$.

[4] See also [14] for a Lagrangian approach in the presence of estimation errors.

by substituting (8a) in the time derivative of (8b).

In this learning context (related itself to the inverse logit choice model of [18]), the auxiliary "score" variable $y_{k\alpha} \in \mathbb{R}$ simply measures how well the eigenvalue $p_{k\alpha}$ has "learned" the gradient vector v which leads to higher sum rates in the unconstrained y-space. Hence, the only difference between the replicator dynamics and exponential learning is one of viewpoint: (7) is written *directly* on the state space of the system (so extra care must be taken in order to satisfy the problem's constraints),[5] while the otherwise equivalent scheme of (8a) is an *unconstrained* descent which relies on the Gibbs distribution (8b) to map solutions back to the system's original state space.

B. Exponential learning in the full matrix problem

Motivated by the above, we introduce the following exponential learning method for the matrix program (MP):

$$\dot{\mathbf{Y}}_k = \mathbf{V}_k, \tag{9a}$$

$$\mathbf{P}_k = P_k \frac{\exp(\mathbf{Y}_k)}{\operatorname{tr}(\exp(\mathbf{Y}_k))}, \tag{9b}$$

where $\mathbf{V}_k \equiv -\partial F/\partial \mathbf{P}_k^*$ is the conjugate derivative of F w.r.t. $\mathbf{P}_k$.[6]

Needless to say, in a MIMO setting, this scheme hinges on users being able to calculate the gradient matrices $\mathbf{V}_k$. To that end, a differentiation of the static sum rate Φ of (2) gives:

$$\mathbf{V}_k = -\frac{\partial F}{\partial \mathbf{P}_k^*} = \frac{\partial \Phi}{\partial \mathbf{P}_k^*} = \mathbf{H}_k^\dagger \mathbf{W}^{-1} \mathbf{H}_k, \tag{10}$$

where $\mathbf{W} = \mathbf{I} + \sum_\ell \mathbf{H}_\ell \mathbf{P}_\ell \mathbf{H}_\ell^\dagger$ is the aggregate signal-plus-noise covariance matrix, assumed to be measured at the receiver end and then made known to the transmitters (e.g. by broadcasting) under the same hypotheses used in standard water-filling schemes [4], [5]. We thus obtain:

Algorithm 1 Exponential Learning

Require: For all $k \in \mathcal{K}$, pick Hermitian initial score matrices $\mathbf{Y}_k \in \mathbb{C}^{m_k \times m_k}$ and positive learning rates $\lambda_k > 0$.
  $t \leftarrow 0$;
  repeat
    $t \leftarrow t + 1$;
    for all $k \in \mathcal{K}$ **do**
      $\mathbf{Y}_k \leftarrow \mathbf{Y}_k + \delta(t)\mathbf{V}_k$;
      $\mathbf{P}_k \leftarrow \exp(\lambda_k \mathbf{Y}_k) / \operatorname{tr}(\exp(\lambda_k \mathbf{Y}_k))$;
    end for
  until required accuracy is reached or transmission ends.

Remark 1. We first note that exponential learning as defined above has the following desirable properties:

(P1) It is *distributed*: users may update the algorithm based on local measurements and information.
(P2) It is *reinforcing*: users move along a direction which increases their individual rates (see also Proposition 3).
(P3) It is *stateless*: users are oblivious to the state of the algorithm, even to the existence of other users.

[5]Note that constraint satisfaction is built into (7) by virtue of the fact that $\sum_{\beta=1}^{m_k} dp_{k\beta}/dt = 0$ and that $dp_{k\beta}/dt = 0$ when $p_{k\beta} = 0$.

[6]Namely, if $P_{\alpha\beta} = X_{\alpha\beta} + iY_{\alpha\beta}$, then the elements of $\partial F/\partial \mathbf{P}^*$ are $\frac{\partial F}{\partial X_{\alpha\beta}} + i\frac{\partial F}{\partial Y_{\alpha\beta}}$.

Remark 2. We also see that *exponential learning does not differentiate between the static and fast-fading regime*: the matrices $\mathbf{H}_k$ (and obviously $\mathbf{W}$) are simply the ones measured at the t-th iteration of the algorithm. In the next section, we will show that if the channel is static, exponential learning converges to the maximum of the static sum rate Φ; otherwise, if the channel matrices evolve ergodically over time, the users converge to the maximum of the ergodic sum rate $\overline{\Phi}$.

Remark 3. The discrete-time implementation of the exponential learning dynamics has two additional (and important) components that are not present in (9): *i*) the users "learning rates" $\lambda_k > 0$; and *ii*) the step sequence $\delta(t)$. The time-step sequence $\delta(t)$ is a standard feature of both deterministic [19] and stochastic [20] optimization algorithms and its role is to ensure convergence when passing from the continuous to the discrete. On the other hand, the role of the learning rate parameter λ_k is much more interesting. Indeed, λ_k can be interpreted as the inverse temperature of the (matrix-valued) Gibbs distribution (9b), and as in simulated annealing, it controls the algorithm's speed: for small λ, learning is slower and smoother, while for larger λ, the algorithm induces rapid changes and then freezes (see also Fig. 1).

IV. Analysis and Convergence Properties

In this section, we will focus on the behavior and convergence properties of the exponential learning dynamics (9) and the corresponding discrete-time algorithm. We thus begin by establishing that the dynamics (9) are *consistent*, i.e. they respect the structure of the matrix state space $\mathcal{X}$:

Proposition 1. *For any Hermitian initialization $\mathbf{Y}_k(0)$, $k \in \mathcal{K}$, the corresponding solution $\mathbf{P}(t)$ of the exponential learning dynamics (9) remains in $\mathcal{X}$ for all t; in particular, $\mathbf{P}_k(t) \succcurlyeq 0$ and $\operatorname{tr}(\mathbf{P}_k(t)) = P_k$ for all $k \in \mathcal{K}$ and for all $t \geq 0$.*

Sketch of proof: Note that $\mathbf{V}_k$ is Hermitian whenever the $\mathbf{Y}_k$ are, so if we start with Hermitian initial conditions in (9a), $\mathbf{Y}_k(t)$ will remain Hermitian for all time. As a result, $\mathbf{P}_k(t) \propto \exp(\mathbf{Y}_k(t))$ will be positive-definite as well, and the trace condition $\operatorname{tr}(\mathbf{P}_k(t)) = P_k$ follows immediately from (9b). ■

Proposition 1 allows us to overcome the important hurdle of consistence in a surprisingly painless way: instead of specifying the dynamics *directly* on $\mathcal{X}$ (a very hard task given the implicit nature of the positivity constraints $\mathbf{P}_k \succcurlyeq 0$), the scheme (9) evolves in an unconstrained space and trajectories are then mapped to the original state space via the Gibbs map (9b).

That said, it is only natural to ask what are the dynamics that govern the evolution of $\mathbf{P}(t)$ in $\mathcal{X}$; to that end, we have:

Proposition 2. *Let $\mathbf{P}(t)$ be an interior solution orbit of the dynamics (9). If $\{p_{k\alpha}(t), \mathbf{u}_{k\alpha}(t)\}$, $\alpha = 1, \dots, m_k$, is an eigensystem of $\mathbf{P}_k$, $k \in \mathcal{K}$, and we set $V_{\alpha\beta}^k \equiv \mathbf{u}_{k\alpha}^\dagger \mathbf{V}_k \mathbf{u}_{k\beta}$, then:*

$$\dot{p}_{k\alpha} = p_{k\alpha}\left(V_{\alpha\alpha}^k - P_k^{-1} \textstyle\sum_{\beta=1}^{m_k} p_{k\beta} V_{\beta\beta}^k\right), \tag{11a}$$

$$\dot{\mathbf{u}}_{k\alpha} = \textstyle\sum_{\beta\neq\alpha} V_{\beta\alpha}^k \left(\log p_{k\alpha} - \log p_{k\beta}\right)^{-1} \mathbf{u}_{k\beta}. \tag{11b}$$

From a mathematical viewpoint, Proposition 2 (proven by taking the Fréchet derivative of (9b) in an eigen-decomposition of $\mathbf{P}$) is a reformulation of the exponential learning dynamics

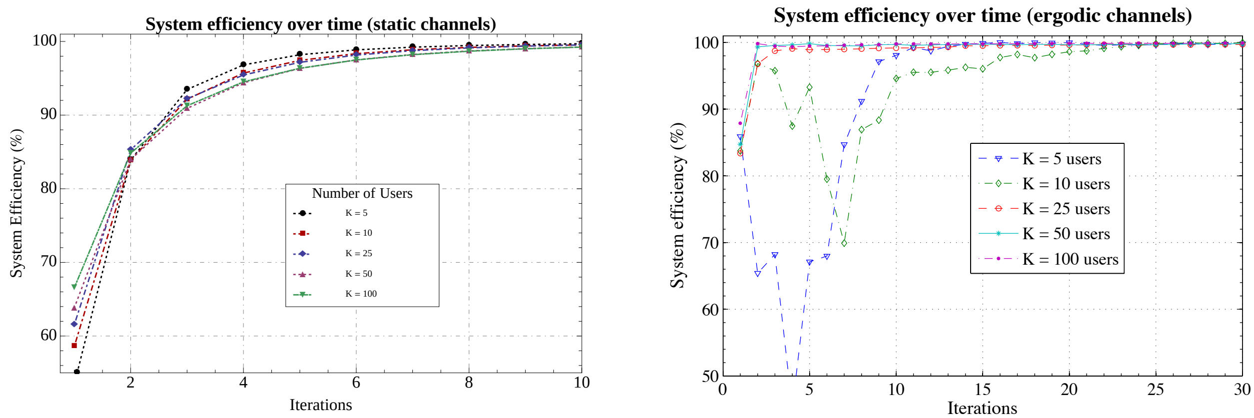

Fig. 1. Convergence of exponential learning (with a constant step size) in static (left) and ergodic channels (right). We plotted the efficiency ratio $\operatorname{eff}(t) = (F(t) - F_{\min})/(F_{\max} - F_{\min})$ for a MIMO MAC system with $n = 5$ receiver antennas and $K = 5, 10, 25, 50$ and 100 users (for simplicity, we only considered diagonal allocations in the ergodic case). In all cases, the system equilibrates rapidly and the convergence rate scales well with the number of users.

(9), so its importance lies in that it illuminates the evolution of the actual state variables $\mathbf{P}_k$.[7] On the other hand, from a computational standpoint, (11) represents a significant simplification of the exponential learning algorithm because users can employ it to update their power allocation policies *directly*, and *without first having to diagonalize/exponentiate the score matrices* $\mathbf{Y}_k$.

This computational benefit is obviously key when one needs to operate with sizable antenna arrays; putting such computational issues aside however, the key property of exponential learning is that its trajectories always descend the objective F:

Proposition 3. *If $\mathbf{P}(t)$ is a (non-stationary) solution of the exponential learning dynamics (9), then $F(\mathbf{P}(t))$ is decreasing.*

Sketch of proof: By the chain rule for matrices, we obtain $dF/dt = \sum_k \operatorname{tr}(\partial F/\partial \mathbf{P}_k^* \cdot \dot{\mathbf{P}}_k^\dagger) = \sum_k \operatorname{tr}(\mathbf{V}_k \dot{\mathbf{P}}_k)$. It can then be shown that this quantity is negative unless $\dot{\mathbf{P}}_k = 0$ (in which case dF/dt vanishes), so our assertion follows. ■

Proposition 3 shows that the system's sum rate always increases along the trajectories of exponential learning; hence, with F convex, we immediately see that *exponential learning converges to the minimum set of F* (i.e. attains its sum capacity). Nonetheless, F need not be *strictly* convex (and in the static case it isn't [7]), so this convergence result does *not* imply that exponentially learning actually converges to a *point*.

From a systems point of view, this is an important question because it is crucial to predict the users' end power allocation policies (and not only the system's sum capacity). To that end, our next result is that exponential learning *does* converge to a point, and, in fact, it converges *exponentially fast*:

Theorem 1. *For any initial Hermitian initialization $\mathbf{Y}_k(0)$, $k \in \mathcal{K}$, the exponential learning dynamics (9) converge to a (possibly initialization-dependent) point $\mathbf{Q}^*$ which minimizes F. Moreover, there exists a positive constant $c > 0$ such that:*

$$\|\mathbf{P}(t) - \mathbf{Q}^*\| \leq \|\mathbf{P}(0) - \mathbf{Q}^*\| \, e^{-ct}. \tag{12}$$

Sketch of proof: The basic ingredient for our proof is the quantum-theoretic generalization of the Kullback-Leibler divergence,[8] which, for $\mathbf{P}, \mathbf{Q} \in \mathcal{X}$, is defined as:

$$D_{\mathrm{KL}}(\mathbf{Q} \,\|\, \mathbf{P}) \equiv \sum\nolimits_k \operatorname{tr}[\mathbf{Q}_k(\log \mathbf{Q}_k - \log \mathbf{P}_k)]. \tag{13}$$

By Klein's inequality [21], we have that $D_{\mathrm{KL}}(\mathbf{Q} \,\|\, \mathbf{P})$ is strictly convex in $\mathbf{P}$ and positive, except at $\mathbf{Q}$ where it vanishes; more importantly, a differentiation of (13) yields the key expression:

$$\frac{d}{dt} D_{\mathrm{KL}}(\mathbf{Q} \,\|\, \mathbf{P}(t)) = -\sum\nolimits_k \operatorname{tr}[(\mathbf{Q}_k - \mathbf{P}_k)\,\mathbf{V}_k], \tag{14}$$

which shows that $\frac{d}{dt} D_{\mathrm{KL}}(\mathbf{Q} \,\|\, \mathbf{P}(t)) \leq 0$ if $\mathbf{Q}$ is a minimum point of F (recall that F is convex so $dF(\mathbf{Q}) \cdot \mathbf{Z} \geq 0$ for all $\mathbf{Z}$ tangent to $\mathcal{X}$ at a minimum point $\mathbf{Q}$).[9] Thus, given that $\mathbf{P}(t)$ converges to the minimum set $\mathcal{X}^*$ of F by Proposition 3, it follows (by compactness of $\mathcal{X}$) that the orbit $\mathbf{P}(t)$ will have an ω-limit $\mathbf{Q}^* \in \mathcal{X}^*$, i.e. $\mathbf{P}(t_n) \to \mathbf{Q}^*$ for some increasing sequence of times $\{t_n\}$, $t_n \to \infty$. Consequently, we will also have $D_{\mathrm{KL}}(\mathbf{Q}^* \,\|\, \mathbf{P}(t_n)) \to 0$, and since the function $D_{\mathrm{KL}}(\mathbf{Q}^* \,\|\, \mathbf{P}(t))$ is itself decreasing, we obtain $\mathbf{P}(t) \to \mathbf{Q}^*$, which proves that $\mathbf{P}(t)$ converges to a point. The convergence rate (12) can then be proven as in [11] (which essentially covers the diagonal case), the details being omitted for lack of space. ■

The above theorem ensures that the continuous exponential learning dynamics (9) always converge to a (global) minimum point of the program's objective (MP). Thus, specializing to the discretized version of the exponential learning algorithm and the MIMO sum rates (2) and (3), we obtain:

Theorem 2. *Let $\mathbf{Y}_k(0)$, $k \in \mathcal{K}$, be a Hermitian initialization of the exponential learning algorithm with time-steps $\delta(t)$ such that $\sum_t \delta(t) = \infty$ and $\sum_t \delta^2(t) < \infty$. Then:*

1) *In static channels, users converge to a power allocation profile $\mathbf{Q}$ which maximizes the sum rate (2).*
2) *In ergodic (fast-fading) channels, users converge (a.s.) to the profile $\mathbf{Q}$ which maximizes their ergodic sum rate (3).*

Sketch of proof: The static case is an Euler approximation scheme with vanishing step size and is thus trivial to dispatch. As for the fast fading regime, recall that exponential learning can be viewed as a dynamical system on $\mathcal{X}$, evolving according

[7] Importantly, if $\mathbf{P}_k$ is diagonal, then the dynamics (11) reduce to the ordinary replicator dynamics – simply compare (11a) to (7). Note also that if $p_{k\alpha} \to 0$, then $\dot{\mathbf{u}}_{k\alpha} \to 0$, so $\mathbf{u}_{k\alpha}$ decouples from the other eigenvectors.

[8] Note that in the diagonal PMAC case, $\mathcal{X}$ is just a product of simplices so power allocation matrices can be interpreted as probability distributions.

[9] Interestingly, this is an alternative proof that $\mathbf{P}(t)$ converges to $\arg\min_{\mathcal{X}} F$.

to the dynamics (11) of Proposition 2. In particular, if the channel matrices $\mathbf{H}_k = \mathbf{H}_k(t)$ are time-dependent (but uncorrelated over time), the eigenvalue dynamics (11a) can be written in discrete time as:

$$p_{k\alpha}(t+1) = p_{k\alpha}(t) + \delta(t) p_{k\alpha}(t) \left[V^k_{\alpha\alpha}(t) - P_k^{-1} \sum\nolimits_{\beta=1}^{m_k} p_{k\beta}(t) V^k_{\beta\beta}(t) \right],$$

where the $V^k_{\alpha\alpha}$ depend on t through both $\mathbf{P}(t)$ and $\mathbf{H}(t)$. Thus, if we set $\overline{V}^k_{\alpha\beta} = \mathbb{E}\left[V^k_{\alpha\beta}\right]$ and $R^k_{\alpha\beta} = V^k_{\alpha\beta} - \overline{V}^k_{\alpha\beta}$, we will have:

$$\begin{aligned} p_{k\alpha}(t+1) &= p_{k\alpha}(t) + \delta(t) p_{k\alpha}(t) \Big[\overline{V}^k_{\alpha\alpha}(t) - P_k^{-1} \sum\nolimits_{\beta=1}^{m_k} p_{k\beta}(t) \overline{V}^k_{\beta\beta}(t) \Big] \\ &+ \delta(t) p_{k\alpha}(t) \left[R^k_{\alpha\alpha}(t) - P_k^{-1} \sum\nolimits_{\beta=1}^{m_k} p_{k\beta}(t) R^k_{\beta\beta}(t) \right], \end{aligned} \tag{15}$$

and similarly for the eigenvector dynamics (11b). Then, by interchanging expectation with differentiation, the resulting matrix $\overline{\mathbf{V}}_k = \{V^k_{\alpha\beta}\}$ of the previous equation corresponds precisely to (minus) the gradient $\partial\overline{\Phi}/\partial\mathbf{P}^*_k$ of the ergodic sum rate $\overline{\Phi}$. Hence, with $\mathbf{H}_k$ stationary and ergodic, the general theory of stochastic approximation (see e.g. Thm. 2 in Chap. 2 of [20]) shows that (15) will track the mean dynamics:

$$\dot{p}_{k\alpha} = p_{k\alpha} \left(\overline{V}_{\alpha\alpha} - P_k^{-1} \sum\nolimits_{\beta=1}^{m_k} p_{k\beta} \overline{V}^k_{\beta\beta} \right), \tag{16}$$

and similarly for the eigenvectors of $\mathbf{P}$. Consequently, since (16) converges to the (globally) minimum point of $\overline{\Phi}$ by Theorem 1, the exponential algorithm will also converge there (a.s.). ■

Remark. A diminishing step size ensures the convergence of the algorithm, but at the expense of convergence speed. With a bit more work however (which we reserve for the future due to space limitations), it can be shown that exponential learning with a constant step *still* converges, always in static channels and with high probability in ergodic ones (see also Fig. 1).

V. Conclusions and Future Work

In this paper, we analyzed the power allocation problem in MIMO multiple access channels by means of an approach based on "exponential learning", a distributed optimization method which applies to general nonlinear problems defined over sets of positive-definite matrices (where traditional optimization methods do not apply because of the form of the problem's constraints). Focusing on the case at hand, we showed that if the system's channels are static, then exponential learning converges to a power allocation profile which attains the sum capacity of the channel exponentially fast; otherwise, if the channels fluctuate stochastically over time following a stationary ergodic process, users converge to a power profile which maximizes their *ergodic* sum rate instead. Importantly, the algorithm's speed can be controlled by tuning the users' learning rate; as a result, the algorithm converges within a few iterations, even when the number of users and/or antennas per user in the system is very large.

Since the method of exponential learning is a rather general one, future applications include the interference channel framework of [5] where the method can be used to attain the Nash equilibria of the channel (in both static and ergodic channels). Moreover, it can also be shown that via the Gibbs transform (9b), exponential learning induces a Riemannian structure on the space of positive-definite matrices, and this structure can be further exploited to yield other learning algorithms, possibly with even faster convergence rates.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications and Related Technologies*, vol. 10, no. 6, pp. 585–596, 1999.

[3] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[4] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, 2004.

[5] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[6] ——, "Competitive design of multiuser MIMO systems based on game theory: a unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, September 2008.

[7] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Dynamic power allocation games in parallel multiple access channels," in *ValueTools '11: ACM Proceedings of the 5th International Conference on Performance Evaluation Methodologies and Tools*, 2011.

[8] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information*, vol. 1, no. 1, pp. 1–182, 2004.

[9] F. Dupuy and P. Loubaton, "On the capacity achieving covariance matrix for frequency selective MIMO channels using the asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5737–5753, September 2011.

[10] K. Tsuda, G. Rätsch, and M. K. Warmuth, "Matrix exponentiated gradient updates for on-line Bregman projection," *Journal of Machine Learning Research*, vol. 6, pp. 995–1018, 2005.

[11] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Distributed learning policies for power allication in multiple access channels," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 96–106, January 2012.

[12] A. Soysal and S. Ulukus, "Optimality of beamforming in fading MIMO multiple access channels," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1171–1183, April 2009.

[13] E. V. Belmega, S. Lasaulce, M. Debbah, and A. Hjørungnes, "Distributed power allocation policies in MIMO channels," in *EUSIPCO '10: Proceedings of the 2010 European Signal Processing Conference*, 2010.

[14] X. Lin and T.-M. Lok, "Learning equilibrium play for stochastic parallel Gaussian interference channels," *submitted*. [Online]. Available: http://arxiv.org/abs/1103.3782v1

[15] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1-2, pp. 145–156, 1978.

[16] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, ser. Economic learning and social evolution. Cambridge, MA: MIT Press, 2011.

[17] P. Mertikopoulos and A. L. Moustakas, "Learning in the presence of noise," in *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*, 2009.

[18] L. E. Blume, "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, vol. 5, pp. 387–424, 1993.

[19] J. A. Snyman, *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*, ser. Applied Optimization. Springer, 2005, vol. 97.

[20] V. S. Borkar, *Stochastic approximation*. Cambridge University Press and Hindustan Book Agency, 2008.

[21] V. Vedral, "The role of relative entropy in quantum information theory," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 197–234, 2002.