BIO-INSPIRED SWARMING FOR DYNAMIC RADIO ACCESS BASED ON DIFFUSION ADAPTATION

†*Paolo Di Lorenzo*, †*Sergio Barbarossa and* ‡*Ali H. Sayed*

†Sapienza University of Rome, DIET, Via Eudossiana 18, 00184 Rome, Italy
‡Electrical Engineering Department, University of California, Los Angeles, CA 90095
E-mail: {dilorenzo,sergio}@infocom.uniroma1.it, sayed@ee.ucla.edu

ABSTRACT

The goal of this paper is to study the learning abilities of adaptive networks in the context of cognitive radio networks and to investigate how well they assist in allocating power and communications resources in the frequency domain. The allocation mechanism is based on a social foraging swarm model that lets every node allocate its resources (power/bits) in the frequency regions where the interference is at a minimum while avoiding collisions with other nodes. We employ adaptive diffusion techniques to estimate the interference profile in a cooperative manner and to guide the motion of the swarm individuals in the resource domain. A mean square performance analysis of the proposed strategy is provided and confirmed by simulation results. Numerical examples show that cooperative spectrum sensing improves the performance of the swarm-based resource allocation technique considerably.

1. INTRODUCTION

Some of the features that are attracting more and more attention in the current research on radio networks are dynamic access, to improve the efficiency of conventional spectrum access protocols [1], and self-organization (SO) capabilities. SO is especially important in femtocell networks, where the deployment of a potentially huge number of user-operated femto-access points makes centralized schemes hard to implement and prone to a heavy signaling traffic. Conversely, decentralized resource allocation strategies are certainly more appealing. However, purely decentralized approaches might lead to highly inefficient systems. Hence, a more viable approach consists in endowing the radio nodes with the capability to learn from the environment and exchange information only with their immediate neighbors, to find out the most appropriate radio resources. Similarly, in cognitive radio, opportunistic (or secondary) users (SU's), are allowed to use temporally unoccupied communication resources, such as frequency bands, time slots or user codes, under the constraint of not interfering (or producing a tolerable interference) with licensed (or primary) users.

The deployment of decentralized radio access strategies was proposed, for example, in [2], where a distributed cooperative spectrum sensing technique exploited the intrinsic sparsity of the radio resource allocation. An important source of inspiration for intrinsically self-organizing systems comes from biological systems, where there is plenty of examples of robust systems, capable of solving difficult organization tasks by exploiting local cooperation among individuals, without the need for a central processor. Recent works illustrate how cooperation over adaptive networks can model collective animal behavior and self-organization in biological networks such as birds flying in formation [3], fish foraging for food [4], or bacteria motility [5]. The application of swarming mechanisms to decentralized resource allocation problems was proposed in [6], where the resource allocation mechanism mimicked the motion of a flock of birds searching for food, assuming a static interference profile and ideal communication channels among swarm individuals. The extension to random interactions among individuals was carried out in [7], to accommodate for random packet drops and quantization noise. In parallel, in [3] it was proposed an adaptive mechanism, based on diffusion adaptation schemes [8,9], to guide the motion of the swarm individuals in *dynamic* environments, through local cooperation. In comparison with other distributed approaches that rely on, for example, consensus-based techniques [10], [11], [12], adaptive networks avoid the need to iterate over data and do not require all nodes to converge to the same equilibrium (or consensus state). Instead, both time- and spatial-diversity of the data are exploited to endow the networks with learning and tracking abilities.

In this paper we combine the swarm-based resource allocation scheme with the diffusion adaptation algorithm. The basic contributions of this paper are: (a) the extension of the social foraging model proposed in [6] to incorporate a real-time distributed spectrum estimation technique based on diffusion adaptation; (b) the derivation of the mean square properties of the diffusion adaptive filter applied to the spectrum estimation problem; and (c) the application of the proposed procedure to the dynamic resource allocation problem in the frequency domain. **Notation:** we use bold face letters to denote random variables and normal font letters to denote their realizations. Matrices and vectors are respectively denoted by capital and small letters.

2. SWARM MODEL

We consider a set of M secondary users who are interested in sharing communications resources in an n-dimensional Euclidean space. A typical setting is one where the resource space is the time-frequency domain (i.e., $n = 2$) and every secondary user is trying to access time and/or frequency slots that are vacant. To keep the notation general, the resource selected by agent k is described by a vector $x_k \in \mathbb{R}^n$, denoting, for example, a frequency subchannel and a time slot. The interaction between the cognitive nodes can be modeled as an undirected graph $G = (V, E)$, where $V \equiv \{1, 2, ..., M\}$ denotes the set of nodes and $E \subseteq V \times V$ is the edge set. Typically, there is a link (edge) between two nodes if the received power from one node to the other node exceeds a minimum threshold value (and this depends on the channel properties). The graph modeling the network topology can be described by the adjacency matrix $A := \{a_{kl}\}$, composed of nonnegative entries $a_{kl} \geq 0$, the degree diagonal matrix D, whose diagonal entries are $d_{kk} := \sum_{l=1}^{M} a_{kl}$, and the Laplacian L, defined as $L = D - A$. We denote by $\mathcal{N}_k$ the set of neighbors of agent k, namely, $\mathcal{N}_k = \{l \in V : a_{kl} \neq 0\}$. We formulate the dynamic radio access problem as the minimization of the

This work has been funded by FREEDOM Project, Nr. ICT-248891. The work of A.H. Sayed was suported in part by NSF grants CCF-1011918 and CCF-0942936.

following potential function

$$J(x,t)=\sum_{k=1}^{M}\sigma_k(x_k,t)+\frac{1}{2}\sum_{k=1}^{M}\sum_{l=1}^{M}a_{kl}J_{ar}(\|x_l-x_k\|),\quad (1)$$

where $\sigma_k(\cdot,\cdot)\in\mathbb{C}^1:\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ represents the interference power over the optimization domain (e.g., the time-frequency plane) perceived by node k at time t and $x:=(x_1^T,\ldots,x_M^T)^T$. The goal of the optimization problem is to find optimal positions x such that the potential function in (1), evaluated a time t, is minimized. The minimization of the first term of (1) leads every node to find a position x_k such that the overall interference power is at a minimum. The second term of (1) is an attraction/repulsion potential function given by

$$J_{ar}(\|x_l-x_k\|)=J_a(\|x_l-x_k\|)-J_r(\|x_l-x_k\|).\quad (2)$$

This potential incorporates a short range repulsion term $J_r(\|x_l-x_k\|)$, whose effect is to avoid collisions among the cognitive nodes, and a long range attraction term $J_a(\|x_l-x_k\|)$, whose goal is to induce a swarm cohesion behavior, e.g. to avoid an excessive spread of the selected radio resources in the time-frequency domain. Hence, in summary, minimizing (1) leads each node to dynamically allocate its own resources in time-frequency regions where there is less interference, helps to avoid conflicts among users, and limits the spread of the occupied domain. We choose attraction and repulsion forces so that there is a unique distance at which the two forces balance: the so called equilibrium distance in the biological literature [13].

A possible way to achieve the distributed minimization of (1) is to use a gradient based optimization procedure, whereby every node starts from an initial guess, say x_0, and then updates its resource allocation vector x_k in time according to the following discrete-time implementation:

$$\begin{aligned}x_k[i+1]&=x_k[i]-\epsilon\nabla_{x_k}J(x,iT)\\&=x_k[i]-\epsilon\nabla_{x_k}\sigma_k(x_k,iT)+\epsilon\textstyle\sum_{l=1}^{M}a_{kl}\;g(x_l[i]-x_k[i]),\end{aligned}\quad (3)$$

$k=1,\ldots,M$, where T is the sampling time, i is the time index, $x(0)=x_0$ and $g(\cdot)$ is a vector function defined as

$$g(x_l-x_k)=[g_a(\|x_l-x_k\|)-g_r(\|x_l-x_k\|)](x_l-x_k),\quad (4)$$

where $g_a(\|x_l-x_k\|)(x_l-x_k)$ and $g_r(\|x_l-x_k\|)(x_l-x_k)$ are the gradients of $J_a(\|x_l-x_k\|)$ and $J_r(\|x_l-x_k\|)$ with respect to x_k, respectively. In this paper we consider attraction/repulsion functions having a constant attraction term of the form $g_a(\|y\|)=c_A$ and an unbounded repulsion term of the form $g_r(\|y\|)=c_R/\|y\|^2$. Unbounded repulsion prevents collisions among the secondary users. The equilibrium distance between attraction and repulsion forces is properly adjusted through the positive parameters c_A and c_R. In our setting, the equilibrium distance is chosen to be proportional to the bandwidth of the frequency subchannel, in the frequency domain, or to the duration of the elementary time slot. Furthermore, the coefficients a_{kl} depend on the distance between the nodes, and two nodes communicates with each other only if they are neighbors. Hence, two nodes k and l with no direct link between them (i.e., for which $a_{kl}=0$), may end up with the same allocation vector; this occurrence is what is known as spatial reuse of frequency or time slots. Under these conditions, it is possible to establish that the swarming algorithm converges exponentially with a speed depending on the profile features and on the second eigenvalue of the Laplacian matrix of the graph.

In the next section, we show how to adaptively estimate the gradient of the interference profile, $\nabla_{x_k}\sigma_k(x_k,t)$, in a distributed manner and through local cooperation. This gradient vector is needed in (3) to update the swarming behavior.

3. DISTRIBUTED SPECTRUM ESTIMATION

3.1 PSD Basis Expansion

Let $\Phi_q(f)$ denote the power spectral density (PSD) of the signal transmitted by the q-th primary user (PU). The PSD can be represented as a linear combination of some preset basis functions, say, as:

$$\Phi_q(f)=\sum_{j=1}^{J}b_j(f)w_{qj}=b_0^T(f)w_q\quad (5)$$

where $b_0(f)=[b_1(f),...,b_J(f)]^T$ is the vector of basis functions evaluated at frequency f, $w_q=[w_{q1},...,w_{qJ}]$ is a vector of weighting coefficients representing the power transmitted by the q-th PU over each basis, and J is the number of basis functions. For J sufficiently large, the basis expansion in (5) can well approximate the transmitted spectrum. In particular, we consider continuously differentiable basis functions, such as raised cosines, Gaussian pulses, etc. Assuming N active users are transmitting, the overall transmitted spectrum can be expressed as

$$\sigma_T(f)=\sum_{q=1}^{N}\sum_{j=1}^{J}b_j(f)w_{qj}=b_1^T(f)w\quad (6)$$

where $b_1(f)=1\otimes b_0(f)$, with $\otimes$ denoting the Kronecker product, and $w=[w_1,\ldots,w_N]\in\mathbb{R}^{JN}$. The propagation medium introduces path loss attenuation between primary and secondary users. Let p_{qk} be the path loss coefficient between the q-th primary user (PU) transmitter and the k-th secondary user (SU). Under the assumption of spatial uncorrelatedness of the channels, the signal received by the secondary node k can be expressed as

$$\sigma_k(f)=\sum_{q=1}^{N}p_{qk}\sum_{j=1}^{J}b_j(f)w_{qj}+\sigma_k^2=b_k^T(f)w+\sigma_k^2\quad (7)$$

where $p_k=[p_{1k},...,p_{Nk}]$ is the vector of path-loss coefficients between every transmitter and the k-th receiver, $b_k(f)=p_k\otimes b_0(f)$ and σ_k^2 is the noise power at the k-th node.

3.2 Diffusion Adaptation

At every time instant i, every node k observes the received PSD in (7) over N_c frequency samples $f_m=f_{min}:(f_{max}-f_{min})/(N_c-1):f_{max}$, for $m=1,\ldots,N_c$, according to the linear model:

$$d_k^m(i)=b_{k,i}^T(f_m)w_0+\sigma_k^2+v_{k,i}^m\quad (8)$$

where $v_{k,i}^m$ is a zero mean random variable with variance $\sigma_{v,m}^2$. The temporal index i in the regressor expression $(b_{k,i}^T(f_m))$ takes into account the possibility of node mobility and possible variations in the channel conditions over time. The receiver noise power σ_k^2 can be pre-estimated with high accuracy using an energy detection over an idle band. It can then be removed from the expression in (8). Collecting measurements over N_c contiguous channels, we obtain a vector linear model:

$$d_k(i)=B_{k,i}w_0+v_{k,i}\quad (9)$$

where $B_{k,i}=[b_{k,i}^T(f_m)]_{m=1}^{N_c}\in\mathbb{R}^{N_c\times JN}$ with $N_c>JN$, and $v_{k,i}$ is a zero mean random vector with covariance matrix $R_{v,i}$. The cooperative estimation problem can be cast as the distributed minimization of the following cost function:

$$J_w(w)=\sum_{k=1}^{M}\mathbb{E}\|\boldsymbol{d}_k(i)-\boldsymbol{B}_{k,i}w\|^2\quad (10)$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator. Several diffusion adaptation schemes have been developed for such purpose in [8]. In this paper, we employ a normalized version of the Adapt-then-Combine (ATC) algorithm without measurement exchange [9]. For the vector minimization problem in (10), the normalized ATC algorithm reads as follows:

$$\begin{cases} \psi_{k,i} = w_{k,i-1} + \mu_i H_{k,i} B_{k,i}^T [d_k(i) - B_{k,i} w_{k,i-1}] \\ w_{k,i} = \sum_{l \in \mathcal{N}_k} c_{l,k} \psi_{l,i} \end{cases} \tag{11}$$

where μ_k is a positive step size chosen by node k and $H_{k,i} = (B_{k,i}^T B_{k,i})^\dagger$, where $(\cdot)^\dagger$ denotes the pseudo-inverse operation. The first step in (11) involves local adaptation where node k updates using the new observations $\{d_i(k), B_{k,i}\}$. The second step is a combination step where the intermediate estimates $\psi_{l,i}$, from the neighborhood $l \in \mathcal{N}_k$, are combined through the coefficients $\{c_{l,k}\}$. The combination matrix $C = \{c_{l,k}\} \in \mathbb{R}^{M \times M}$ satisfies $c_{l,k} \geq 0$ if $l \in \mathcal{N}_k$ and $1^T C = 1^T$. The resulting estimate of node k at time i is denoted by $w_{k,i}$. In the case in which the unknown parameter w_0 varies slowly with time, the ATC diffusion algorithm allows online tracking of the interference profile variations enabling the swarm to dynamically allocate resources over the frequency domain. Then, referring to expression (6), an estimate of the derivative of the interference profile in the frequency f_m at time i can be computed by node k as:

$$\sigma'_{k,i}(f_m) = \sum_{q=1}^{N} \sum_{j=1}^{J} b'_j(f_m) w_{k,i}^{q,j} = b'^T_{k,i}(f_m) w_{k,i} \tag{12}$$

where $b'_j(f)$ is the known derivative of the j-th basis function.

3.3 Performance Analysis

In this section we analyze the performance of the normalized diffusion algorithm following the approach of [9] and extending it to the case of a vector linear model as in (9). In what follows we view the estimates $w_{k,i}$ as realizations of a random process $\boldsymbol{w}_{k,i}$ and analyze the performance of the algorithms in terms of their mean square behavior. We consider a general algorithmic form that includes various normalized diffusion algorithms as special cases. Thus, we consider a general normalized diffusion filter of the form

$$\begin{cases} \boldsymbol{\phi}_{k,i} = \sum_{l=1}^{M} c_{1,l,k} \boldsymbol{w}_{l,i-1} \\ \boldsymbol{\psi}_{k,i} = \boldsymbol{\phi}_{k,i} + \mu_k \sum_{l=1}^{M} s_{l,k} \boldsymbol{H}_{k,i} \boldsymbol{B}_{k,i}^T [\boldsymbol{d}_k(i) - \boldsymbol{B}_{k,i} \boldsymbol{\phi}_{k,i}] \\ \boldsymbol{w}_{k,i} = \sum_{l=1}^{M} c_{2,l,k} \boldsymbol{\phi}_{k,i} \end{cases} \tag{13}$$

where the coefficients $\{c_{1,l,k}\}$,$\{s_{l,k}\}$ and $\{c_{2,l,k}\}$ are generic non-negative real coefficients corresponding to the (l,k) entries of the matrices C_1, S, and C_2, respectively, satisfying

$$1^T C_1 = 1^T, 1^T S = 1^T, 1^T C_2 = 1^T. \tag{14}$$

Different types of algorithms can be obtained as special cases of (13) by choosing different matrices $\{C_1, S, C_2\}$. The ATC diffusion algorithm without measurement exchange in (11) is obtained by choosing $C_1 = S = I$ and $C_2 = C$. To proceed with the analysis, we assume a linear measurement as in (9). Using (13), we define the error quantities $\tilde{\boldsymbol{w}}_{k,i} = w_0 - \boldsymbol{w}_{k,i}$, $\tilde{\boldsymbol{\psi}}_{k,i} = w_0 - \boldsymbol{\psi}_{k,i}$, $\tilde{\boldsymbol{\phi}}_{k,i-1} = w_0 - \boldsymbol{\phi}_{k,i-1}$ and the global vectors:

$$\tilde{\boldsymbol{w}}_i = \begin{bmatrix} \tilde{\boldsymbol{w}}_{1,i} \\ \vdots \\ \tilde{\boldsymbol{w}}_{M,i} \end{bmatrix} \quad \tilde{\boldsymbol{\psi}}_i = \begin{bmatrix} \tilde{\boldsymbol{\psi}}_{1,i} \\ \vdots \\ \tilde{\boldsymbol{\psi}}_{M,i} \end{bmatrix} \quad \tilde{\boldsymbol{\phi}}_{i-1} = \begin{bmatrix} \tilde{\boldsymbol{\phi}}_{1,i-1} \\ \vdots \\ \tilde{\boldsymbol{\phi}}_{M,i-1} \end{bmatrix}. \tag{15}$$

We also introduce a diagonal matrix

$$\mathcal{M} = \text{diag}\{\mu_1 I_{JN}, \ldots, \mu_M I_{JN}\} \tag{16}$$

and the extended weighting matrices

$$\mathcal{C}_1 = C_1 \otimes I_{JN} \quad \mathcal{C}_2 = C_2 \otimes I_{JN} \quad \mathcal{S} = S \otimes I_{JN} \tag{17}$$

where $\otimes$ denotes the Kronecker product operation. We further introduce the following quantities

$$\boldsymbol{D}_i = \text{diag}\left\{ \sum_{l=1}^{M} s_{l,k} \boldsymbol{H}_{l,i} \boldsymbol{B}_{l,i}^T \boldsymbol{B}_{l,i} \right\}_{k=1}^{M} \tag{18}$$

$$\boldsymbol{g}_i = \mathcal{S}^T col\{\boldsymbol{H}_{k,i} \boldsymbol{B}_{k,i}^T \boldsymbol{v}_k(i)\}_{k=1}^{M}. \tag{19}$$

The matrices $\boldsymbol{H}_{k,i}$ and $\boldsymbol{B}_{k,i}$ depend on the path loss vector $p_{k,i}$ of node k at time i. Let $\gamma_i = [p_{k,i}]_{k=1}^{M}$ denote the overall path loss vector of the network at time i. Then, the matrix $\boldsymbol{D}_i$ and the noise vector $\boldsymbol{g}_i$ are function of γ_i and we have

$$\begin{aligned} \tilde{\boldsymbol{\phi}}_{i-1} &= \mathcal{C}_1^T \tilde{\boldsymbol{w}}_{i-1} \\ \tilde{\boldsymbol{\psi}}_i &= \tilde{\boldsymbol{\phi}}_{i-1} - \mathcal{M}[\boldsymbol{D}_i(\gamma_i) \tilde{\boldsymbol{\phi}}_{i-1} + \boldsymbol{g}_i(\gamma_i)] \\ \tilde{\boldsymbol{w}}_i &= \mathcal{C}_2^T \tilde{\boldsymbol{\psi}}_i \end{aligned} \tag{20}$$

or, equivalently,

$$\tilde{\boldsymbol{w}}_i = \mathcal{C}_2^T [I - \mathcal{M} \boldsymbol{D}_i(\gamma_i)] \mathcal{C}_1^T \tilde{\boldsymbol{w}}_{i-1} - \mathcal{C}_2^T \mathcal{M} \boldsymbol{g}_i(\gamma_i). \tag{21}$$

3.3.1 Mean Stability

Assuming the regression data are spatially and temporally white and taking the expectation of (21), we get

$$\mathbb{E}\tilde{\boldsymbol{w}}_i = \mathcal{C}_2^T [I - \mathcal{M} \mathbb{E} \boldsymbol{D}_i(\gamma_i)] \mathcal{C}_1^T \mathbb{E}\tilde{\boldsymbol{w}}_{i-1} \tag{22}$$

Using Lemma 1 from [9], we conclude that $\tilde{\boldsymbol{w}}_i$ is asymptotically unbiased if the matrix $I - \mathcal{M}\mathcal{D}_i(\gamma_i)$, where $\mathcal{D}_i(\gamma_i) = \mathbb{E}\boldsymbol{D}_i(\gamma_i)$, is stable for all i. Then, exploiting the expression in (18), the algorithm converges in the mean for any step-size satisfying $0 < \mu_k < 2$ for all k.

3.3.2 Mean Square Performance

In this section we examine the mean-square performance of the diffusion filter (13). Now, following the energy conservation arguments of [14], we evaluate the weighted norm of $\tilde{\boldsymbol{w}}_i$ obtaining

$$\begin{aligned} \mathbb{E}\|\tilde{\boldsymbol{w}}_i\|_\Sigma^2 &= \mathbb{E}\|\tilde{\boldsymbol{w}}_{i-1}\|^2_{\mathcal{C}_1(I - \mathcal{D}_i(\gamma_i)\mathcal{M})^T \mathcal{C}_2 \Sigma \mathcal{C}_2^T (I - \mathcal{M}\mathcal{D}_i(\gamma_i)) \mathcal{C}_1^T} + \\ &+ \mathbb{E}[\boldsymbol{g}_i^T(\gamma_i) \mathcal{M} \mathcal{C}_2 \Sigma \mathcal{C}_2^T \mathcal{M} \boldsymbol{g}_i(\gamma_i)] \end{aligned} \tag{23}$$

where Σ is an Hermitian positive-definite matrix that we are free to choose. Moreover, setting

$$G(\gamma_i) = \mathbb{E}[\boldsymbol{g}_i(\gamma_i) \boldsymbol{g}_i^T(\gamma_i)] \tag{24}$$

we can rewrite (23) as a variance relation of the form

$$\mathbb{E}\|\tilde{\boldsymbol{w}}_i\|_\Sigma^2 = \mathbb{E}\|\tilde{\boldsymbol{w}}_{i-1}\|_{\Sigma'}^2 + \text{Tr}[\Sigma \mathcal{C}_2^T \mathcal{M} G(\gamma_i) \mathcal{M} \mathcal{C}_2] \tag{25}$$

where $\text{Tr}(\cdot)$ is the trace operator, and

$$\begin{aligned} \Sigma' = \mathcal{C}_1 \mathcal{C}_2 \Sigma \mathcal{C}_2^T \mathcal{C} 1^T - \mathcal{C}_1 \mathcal{D}_i(\gamma_i) \mathcal{M} \mathcal{C}_2 \Sigma \mathcal{C}_2^T \mathcal{C}_1^T + \\ - \mathcal{C}_1 \mathcal{C}_2 \Sigma \mathcal{C}_2^T \mathcal{M} \mathcal{D}_i(\gamma_i) \mathcal{C}_1^T + \mathcal{C}_1 \mathcal{D}_i(\gamma_i) \mathcal{M} \mathcal{C}_2 \Sigma \mathcal{C}_2^T \mathcal{M} \mathcal{D}_i(\gamma_i) \mathcal{C}_1^T. \end{aligned} \tag{26}$$

Let $\sigma = \text{vec}(\Sigma)$ denote the vector that is obtained by stacking the columns of Σ on top of each other. Using the Kronecker product property $\text{vec}(U\Sigma V) = (V^T \otimes U)\text{vec}(\Sigma)$, we can rewrite Σ' in (26) as $\sigma' = \mathrm{vec}(\Sigma') = F\sigma$, where the matrix F is given by

$$\begin{aligned} F &= (\mathcal{C}_1 \otimes \mathcal{C}_1)\{I - I \otimes (\mathcal{D}_i(\gamma_i)\mathcal{M}) - (\mathcal{D}_i(\gamma_i)\mathcal{M}) \otimes I + \\ &+ (\mathcal{D}_i(\gamma_i)\mathcal{M}) \otimes (\mathcal{D}_i(\gamma_i)\mathcal{M})\}(\mathcal{C}_2 \otimes \mathcal{C}_2). \end{aligned} \tag{27}$$

In the following we assume that the path loss vector $\gamma_i \to \gamma_0$, where γ_0 is a fixed constant vector, as $i \to \infty$. Then, using the property $\mathrm{Tr}(\Sigma X) = \mathrm{vec}(X^T)^T\sigma$ and taking the limit of (25) as $i \to \infty$, we can recast (25) as follows:

$$\mathbb{E}\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma-\Sigma'} = [\mathrm{vec}(\mathcal{C}_2^T \mathcal{M} G(\gamma_0)^T \mathcal{M}\mathcal{C}_2)]^T\sigma. \tag{28}$$

The steady-state mean-square deviation (MSD) at node k is defined as:

$$\mathrm{MSD}_k = \lim_{i\to\infty} \mathbb{E}\|\tilde{\boldsymbol{w}}_{k,i}\|^2. \tag{29}$$

Then, if the step sizes $\{\mu_k\}$ are small enough so that the matrix $(I - F)$ is invertible and choosing $\sigma = (I - F)^{-1}m_k$, the MSD of node k tends to

$$\mathrm{MSD}_k = [\mathrm{vec}(\mathcal{C}_2^T \mathcal{M} G(\gamma_0)^T \mathcal{M}\mathcal{C}_2)]^T (I - F)^{-1} m_k, \tag{30}$$

where $m_k = \mathrm{vec}(\mathrm{diag}(e_k) \otimes I_{JN})$, with e_k being column vectors with a unity entry at position k and zeros elsewhere. In the next section we illustrate how these theoretical expressions match well with simulation results.

4. SIMULATION RESULTS

In this section, we provide numerical examples illustrating the performance of the proposed technique combining the swarm-based resource allocation method, illustrated in Section 2, and the distributed cooperative sensing algorithm using ATC diffusion adaptation, shown in Section 3. We consider a connected network composed of 15 SU's, plus two PU's. The topology of the network is shown in Fig. 1, where the green dots represent SU's, while the red dots indicate the PU's. We assume that a PU moves from the initial position, indicated as the orange dot, to the final position represented by the red dot, so that the interference perceived by the secondary network is time-varying. We consider a polynomial path loss model $p_{qk}(d_{qk}) = (d_{qk}/d_0)^{-\gamma}$ ($\gamma = 2$), where d_{qk} is the distance between the q-th PU and the k-th SU. The cognitive SU's scan $N_c = 80$ channels between 30 and 45 MHz and use $J = 16$ Gaussian basis functions to model the basis expansion of the transmitted spectrum. In this simulation, we consider a combination matrix C that simply averages the intermediate estimate from the neighborhood, hence, such that $c_{l,k} = 1/\mathcal{N}_k$ for all l. We assume the presence of 15 resources (to be allocated from as many cognitive users) that are initially scattered randomly across the frequency spectrum. At the k-th iteration of the updating rule (3), each node communicates to its neighbors the position it intends to occupy, i.e., the scalar $x_i[k]$ representing a frequency subchannel. In the application at hand, there is an intrinsic quantization of the frequency resources given by the subchannel bandwidth. In our implementation, we let the system evolve according to (3) until successive differences in allocation become smaller than the bandwidth of a frequency subchannel. At that point, the evolution stops and every SU is allowed to transmit over the selected channel. We consider an interference profile as in Fig. 2, where the dashed black curve depicts the true transmitted spectrum, whereas the continuous red and dashed blue curves represent, respectively, the estimation obtained using ATC diffusion (at convergence) and without cooperation among the nodes. We notice how diffusion adaptation fits well the spectrum profile while the non-cooperative approach yields a rather poor estimation. To evaluate the performance of the distributed estimation technique, in Fig. 3 we show the steady-state MSD of the ATC diffusion algorithm compared with the theoretical result in (30). The steady-state values are obtained by averaging over 200 independent experiments and over 100 time samples after convergence. We can observe that the simulation results match perfectly with the theoretical values. An example of resource allocation is shown in Fig. 2, where the green dots represent the final frequency channels chosen at convergence by the network nodes. The parameters of the swarm are $c_A = 0.025$, $c_R = 0.25$. It is evident how the resources avoid the position occupied by primary users, tend to keep the spread as small as possible while avoiding collisions among the allocations of different users. Observe that the number of allocated channels is less than the number of requested resources. This means that a certain number of nodes have picked up the same channels. Nevertheless, we have checked numerically that the nodes having chosen the same resource are not neighbors of each other, so that there are no real collisions. In other words, the algorithm is capable of implementing a decentralized mechanism for spatial reuse of frequencies. To measure the effectiveness of the distributed resource allocation strategy, in Fig. 4 we report the interference level, versus the number of nodes composing the secondary network, averaged over the frequency slots occupied by the SUs, after convergence. The result is averaged over 200 independent realizations. We considered two different values of the re-

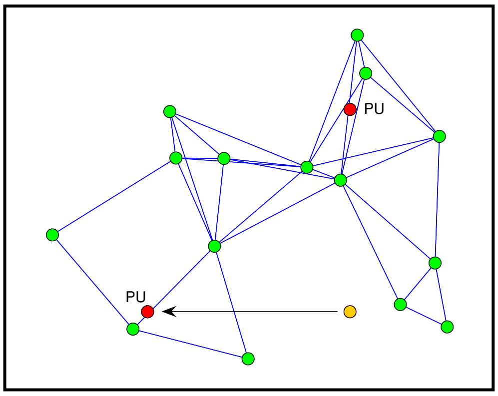

Figure 1: Secondary network. The red nodes denote primary users and the green nodes denote secondary users.

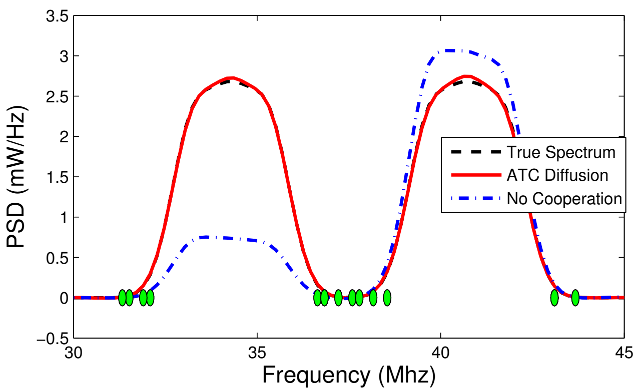

Figure 2: Comparison of the result of spectrum estimation through cooperative diffusion adaptation and without cooperation among the users.

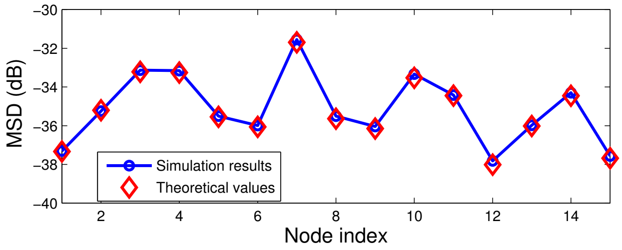

Figure 3: Steady-state MSD versus node index.

ceiver noise power σ^2, which determines the variance of the estimation noise $v_{k,i}^m$ in (8). The parameters of the swarm are $c_A = 0.025$, $c_R = 0.25$ and the interference profile is the same considered in Fig. 2. From Fig. 4, we notice that, using a non cooperative approach, the estimation of the interference profile gradient is quite poor and some resources end up being allocated by mistake in the regions occupied by the primary users, trapped because of the estimation errors affecting the algorithm. This explains the high level of interference perceived in this case. The performance of the allocation can be remarkably improved adopting the cooperative diffusion adaptation approach. Indeed, as the estimation accuracy improves, each resource tends to move towards the interference-free regions, thus making the overall swarm experience a smaller total interference. As the number of nodes M increases, the allocation performance improves because the swarming algorithm exploits a cooperative capability given by the cohesion force. This intrinsic robustness determines that the agents, allocated over the low interference bands, tend to form cohesive blocks that exert an attraction towards the agents trapped by mistake over the regions of the spectrum occupied by the primary users. Moreover, in the cooperative case, an increase in the number of nodes also improves the estimation performance, thus simplifying the resource allocation task. From Fig. 4, we also note, as expected, how a stronger noise leads to worst allocation performance in both cases. Nevertheless, the performance of the cooperative approach is less sensitive. This means that the performance of the resource allocation based on the swarming algorithm can be considerably improved if every node cooperates with its own neighbors to adaptively estimate the interference profile.

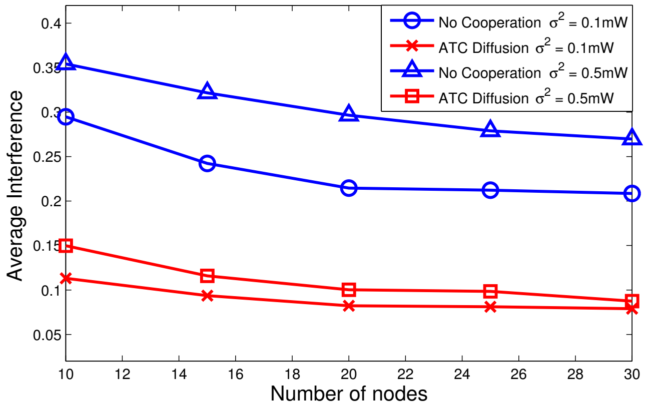

Figure 4: Average interference perceived by the swarm at convergence, for the non cooperative estimation case and for adaptive diffusion.

5. CONCLUSIONS

In this paper we have proposed a dynamic resource allocation technique combining a distributed diffusion algorithm, for implementing cooperative sensing, with a swarming technique, for allocating resources in a parsimonious way (i.e., avoiding unnecessary spread in the frequency domain), yet avoiding collisions. A mean square analysis for the diffusion adaptation filter has been derived and simulation results match well with the theoretical results. Finally, the procedure has been applied to the dynamic resource allocation problem in the frequency domain. Numerical results show the improvement that results in the resource allocation performance due to the cooperative estimation of the spectrum.

REFERENCES

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal. Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[2] J. A. Bazerque, and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol 58, No. 3, pp. 1847–1862, March 2010

[3] F. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2038–2051, May 2011.

[4] S-Y. Tu and A. H. Sayed, "Foraging behavior of fish schools via diffusion adaptation," *Proc. International Workshop on Cognitive Information Processing*, pp. 63–68, Elba Island, Italy, June 2010.

[5] J. Chen, X. Zhao, and A. H. Sayed, "Bacterial motility via diffusion adaptation," Proc. *44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2010.

[6] P. Di Lorenzo, S. Barbarossa, "Distributed resource allocation in cognitive radio systems based on social foraging swarms," in *Proc. 11th IEEE Int. Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, Marrakech, June 2010.

[7] P. Di Lorenzo, S. Barbarossa, "Bio-inspired swarming models for decentralized radio access incorporating random links and quantized communications," in *Proc. 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5780–5783, Prague, May 2011.

[8] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, July 2008.

[9] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. on Signal Processing*, vol. 58, pp. 1035–1048, March 2010.

[10] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus" *Proc. Information Processing in Sensor Networks*, Los Angeles, CA, pp. 63–70, Apr. 2005.

[11] S. Barbarossa, G. Scutari, "Bio-inspired sensor network design: Distributed decision through self-synchronization" *IEEE Signal Processing Magazine*, Volume 24, Issue 3, pp. 26–35, May 2007.

[12] I. D. Schizas, G. B. Giannakis, S. D. Roumeliotis and A. Ribeiro, "Consensus in ad hoc WSNs with noisy links - part II: distributed estimation and smoothing of random signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1650–1666, April 2008.

[13] D. Grnbaum and A. Okubo, "Modeling social animal aggregations" *Frontiers in Theoretical Biology.* New York: Springer-Verlag, 1994,

[14] A. H. Sayed, *Adaptive Filters*, Wiley, NJ, 2008.