Statistical Routing for Multihop Wireless Cognitive Networks

Emiliano Dall'Anese, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

Abstract—To account for the randomness of propagation channels and interference levels in hierarchical spectrum sharing, a novel approach to multihop routing is introduced for cognitive random access networks, whereby packets are randomly routed according to outage probabilities. Leveraging channel and interference level statistics, the resultant cross-layer optimization framework provides optimal routes, transmission probabilities, and transmit-powers, thus enabling cognizant adaptation of routing, medium access, and physical layer parameters to the propagation environment. The associated optimization problem is non-convex, and hence hard to solve in general. Nevertheless, a successive convex approximation approach is adopted to efficiently find a Karush-Kuhn-Tucker solution. Augmented Lagrangian and primal decomposition methods are employed to develop a distributed algorithm, which also lends itself to online implementation. Enticingly, the fresh look advocated here permeates benefits also to conventional multihop wireless networks in the presence of channel uncertainty.

Index Terms—Routing, cross-layer optimization, multihop wireless networks, cognitive radios, random access, channel uncertainty, convex approximation, distributed computation.

I. INTRODUCTION

RESEMBLING traditional routing protocols for wired networks, their counterparts for wireless networking generally utilize optimization tools such as shortest path routing to find optimal route(s) based on the network connectivity graph abstraction [1]. Early on, links among nodes were quantified based on a disk model capturing only distance-based deterministic losses. Upon recognizing the inadequacy of disk models for the broadcast wireless interface [2], a weighted graph accommodating more sophisticated performance metrics was adopted; see e.g., [3], [4], and the stochastic routing approach in [5], where link weights capturing packet delivery probabilities were exploited to develop optimal routing schemes. These schemes are particularly attractive for energy-limited nodes, primarily because the resulting routing strategies promote links with higher reliability, thus decreasing the number of packet lost due to fading [6].

In a hierarchical access setting, interference levels can not be acquired accurately due to the lack of explicit inter-system cooperation [7]. As a result, random shadowing and small-scale fading effects, along with dynamically changing activities of licensed users, accentuate the uncertain nature of wireless cognitive radio (CR) links. The effects of random interference on CR links from primary user (PU) transmitters is called upon in [8], where source-to-destination paths that are most likely to meet prescribed end-to-end requirements are found based on predicted link capacities. Leveraging the situational-awareness provided by spectrum occupancy detection schemes, a graph whose link weights reflect the amount of spectral resources available per CR-to-CR link is employed in [9], where optimal routes are obtained via Dijkstra or Bellman Ford-like algorithms. A two-phase approach is proposed in [10], where nodes in the network first obtain an expected route cost and a set of candidate forwarding nodes, and then route traffic across paths with higher spectrum availability. In [11], the average link availability is invoked to develop a routing scheme that avoids network zones with unstable CR connectivity. Link availability in [11] is computed in a probabilistic sense based only on the statistics of primary user (PU) activities. A PU coverage map supplied by sensing schemes is employed in [12] to identify spectrum opportunities in space, and devise routing strategies supporting multiple classes of CR quality-of-service (QoS) demands.

The aforementioned works offer valuable insights on route formation and management based on the average availability of CR links, and predicted link capacities. However, in a hierarchical access setup, link capacities are unknown and may change abruptly because of time-varying PU activity patterns, dynamic shadowing, and diverse QoS constraints. In this context, a cross-layer design approach to obtain *both* optimal routes *and* physical and medium access parameters that dictate the packet forwarding capabilities is therefore well motivated. To this end, the present paper exploits propagation channel statistics to develop a statistical routing approach whereby nodes not only compute optimal routes, but also optimal link reliabilities by controlling transmit-powers and medium access control (MAC) parameters. The novel approach accounts explicitly for the randomness of propagation and the medium access interface, to allow spectrum-cognizant routing of data packets, while enforcing PU interference protection (Section II).

In spite of the non-convexity of the associated cross-layer optimization problem, a successive convex approximation is pursued to find a Karush-Kuhn-Tucker (KKT) solution efficiently (Section III). Enticingly, feasibility guarantees offered by the successive convex approximation algorithm naturally suggest an online implementation of the algorithm whereby

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA (e-mail: {emiliano, georgios}@umn.edu).

nodes do not necessarily wait for the successive convex approximation iterations to converge, but rather use network parameters as they become available.

However, the communication overhead incurred to acquire channel statistics at a central node, and subsequently disseminate optimal network parameters can become prohibitive as the network size increases. To alleviate such a message-passing burden, and address scalability and robustness concerns, a distributed algorithm is also developed by invoking the alternating direction method of multipliers and the primal decomposition method (Section IV). Finally, suitable conditions are established to ensure that packets are eventually delivered to their destination when routes, medium access and physical layer parameters are regularly updated to track channel statistics and topology dynamics (Section V).

A. Preliminaries and problem formulation

Consider a wireless CR network with N nodes $\{U_n\}_{n=1}^{N}$ sharing spectral resources with an incumbent PU system [7]. Leveraging the spectrum awareness provided by spatio-temporal sensing schemes [13], [14], CRs collaborate in routing data packets to a sink node U_{N+1}, while respecting the PU-CR hierarchy. The CR network is modeled as a digraph to account for the possible lack of link bi-directionality. The dynamic and stochastic nature of the CR propagation ambience, along with the possibly minimal amount of topological information motivate consideration of random medium access, as well as stochastic routing strategies [4], [5], [15]. In this context, a CR node U_n transmits with probability $\mu_n \in [0,1]$, and decides whether to route packets toward a neighboring node U_i with probability $t_{n \to i} \in [0,1]$ per time slot. As packets are forwarded to neighboring nodes according to probability mass functions, it holds that $\sum_{i \neq n} t_{n \to i} = 1$, for all $n = 1, \ldots, N$.

Communication of data packets over a wireless network depends not only on transmission and forwarding decisions, but also on the intended link reliability. In case of unsuccessful packet decoding due to fading- or interference-induced link outages [2], a packet not eventually routed by U_n will remain in U_n's queue, and its transmission will be re-attempted in a subsequent time slot (possibly to a different neighboring CR). To capture channel- and interference-induced sources of uncertainty, let $r_{n \to i} \in (0,1]$ denote the probability that a packet transmitted from node U_n is correctly decoded (and thus successfully received) by U_i.

Assuming that link reliabilities $\{r_{n \to i}\}$ are *known* by, e.g., computing the packed error rate of preceding sessions, a stochastic routing framework for maximizing users' exogenous rates was introduced in [5]. However, because of the volatile CR channel characteristics, time-varying PU activity patterns, and diverse QoS constraints, $\{r_{n \to i}\}$ may change abruptly during the network operation. Hence, $\{r_{n \to i}\}$ may *not* be known in advance. Building on first- and second-order statistics of the PU interference, as well as those of node-to-node channels, a statistical routing approach yielding optimal (i) routes, (ii) transmission probabilities, and (iii) transmit-powers is put forward in the ensuing section.

II. STATISTICAL ROUTING FRAMEWORK

Data percolation through a wireless network is captured by the product packet delivery probabilities $\{t_{n \to i} \, r_{n \to i}\}$. When random access is employed as MAC, it is common to consider a packet lost when collisions among CR transmissions occur. With $\mathcal{I}_{ni}$ denoting the set of nodes whose transmissions interfere with link $U_n \to U_i$, the probability of collision-free packet transmission from U_n to U_i is given by $\prod_{j \in \mathcal{I}_{ni}} (1 - \mu_j)$. A widely-accepted criterion for successful packet reception is to require the signal-to-interference-plus-noise ratio (SINR) to stay above a certain threshold [2], [6], which is generally determined by the receiver structure, modulation, and coding scheme. Let $g_{n \to i}$ denote the channel gain between U_n and U_i, modeling the effects of path loss, log-normal shadowing, and Nakagami-m small-scale fading [16]. Then, the SINR of link $U_n \to U_i$ can be expressed as

$$\gamma_{n \to i} := \frac{p_n g_{n \to i}}{\sigma_i^2 + \sum_{S=1}^{N_S} \pi_{S,i}} \qquad (1)$$

where σ_i^2 stands for the receiver noise power at U_i; $p_n \in (0, p_n^{\max}]$ denotes the transmission power of U_n; and $\pi_{S,i}$ the interference perceived from PU transmitter $S = 1, \ldots, N_S$. Randomness of $\{\gamma_{n \to i}\}$ in (1) emerges due to the shadowing and small-scale effects on the PU interference $\{\pi_{S,i}\}$. Furthermore, CR-to-CR gains $\{g_{n \to i}\}$ may be known imperfectly because of insufficient time for channel training. Nonetheless, CR-to-CR and PU-to-CR deterministic path losses, and statistics of shadowing and small-scale fading can be acquired and used. To this end, it is useful to recall that the distribution of channel gains $\{g_{n \to i}\}$ can be approximated as log-normal [16, Ch. 2], [17]. Furthermore, the Fenton-Wilkinson result [18] asserts that the distribution of SINRs $\{\gamma_{n \to i}\}$ in (1) can be well-approximated as log-normal too, with mean and variance expressed in terms of the first- and second-order moments of $\{g_{n \to i}\}$ and $\{\pi_{S,i}\}$; see [17] for a detailed derivation. Consequently, $\Gamma_{n \to i} := 10 \log_{10} \gamma_{n \to i}$ will be approximately Gaussian distributed with mean $P_n + m_{n \to i}$, where $P_n := 10 \log_{10} p_n$, and variance denoted by $\sigma_{n \to i}^2$. The probability $r_{n \to i}$ that a packet transmitted from U_n is correctly received by U_i can thus be expressed as

$$\begin{aligned} r_{n \to i} &= \prod_{j \in \mathcal{I}_{ni}} (1 - \mu_j) \Pr\{\gamma_{n \to i} > \bar{\gamma}_{n \to i}\} \\ &\approx \prod_{j \in \mathcal{I}_{ni}} (1 - \mu_j) Q\left(\frac{\bar{\Gamma}_{n \to i} - P_n - m_{n \to i}}{\sigma_{n \to i}} \right) \end{aligned} \qquad (2)$$

where $Q(x) := \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx$ is the standard Gaussian tail function, $\bar{\gamma}_{n \to i}$ is a prescribed SINR threshold, and $\bar{\Gamma}_{n \to i} := 10 \log_{10} \bar{\gamma}_{n \to i}$. Similar to [19], the main interest here is in the tail of the complementary cumulative density function (ccdf) of the SINR; in this case, the Fenton-Wilkinson method is known to provide accurate approximations for all the propagation scenarios of practical interest [16, Ch. 3], [17], [20].

Using (2), the link reliabilities $\{r_{n \to i}\}$ can be expressed in terms of the MAC variables $\{\mu_n\}$ and the physical layer quantities $\{P_n, m_{n \to i}, \sigma_{n \to i}\}$. Therefore, with $\{m_{n \to i}, \sigma_{n \to i}\}$ known parameters, the optimal routing strategy will be obtained by optimizing over $\{\mu_n, t_{n \to i}, P_m\}$. The next step is to
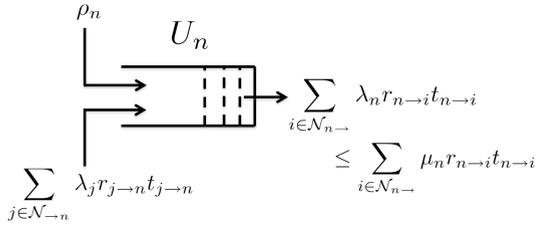
Fig. 1.   Input and output flows at node U_n under queue stability.

model exogenous data packet arrivals at U_n from its application layer by a stationary stochastic process with average rate $\rho_n \in (0,1]$ per time slot. Suppose also that each CR node maintains a backlog to cache exogenous and endogenous[1] packets that have to be routed toward the destination U_{N+1}. Aggregate queue service rates depend on the joint queue occupancy distribution. This results in a generally asymmetric system of interacting queues, whose stability region is challenging to analyze even for simple systems. Nevertheless, assuming as usual fully backlogged queues per node [21] yields a sufficient condition for queue stability that can be conveniently used as a constraint in rate-oriented routing optimization. In the advocated dominant system, users with empty queues transmit "dummy" packets, and consequently queue sizes are never smaller than those in the original system, if both systems start from the same initial condition.

Let λ_n denote the average aggregate rate of endogenous packet arrivals at U_n, which coincides with the rate of packet departures if queues are stable. Then, queue stability implies that $\{\rho_n\}_{n=1}^N$ and $\{\lambda_n\}_{n=1}^N$ abide by the flow conservation constraints [5], [22] (cf. Fig. 1)

$$\rho_n = \lambda_n \sum_{j\in\mathcal{N}_{n\to}} t_{n\to j} r_{n\to j} - \sum_{i\in\mathcal{N}_{\to n}} \lambda_i t_{i\to n} r_{i\to n}, \ \forall n \quad (3)$$

where $\mathcal{N}_{n\to} := \{j | r_{n\to j} > 0, j = 1, \ldots, N+1, j \neq n\}$ is the set of nodes that decode U_n's transmissions with non-zero probability, and $\mathcal{N}_{\to n} := \{i | r_{i\to n} > 0, i = 1, \ldots, N, i \neq n\}$ the set of nodes that route packets through U_n. For queue stability, Loynes' Theorem [23] asserts that for stationary arrival and departure processes (the latter are stationary in the dominant system) a sufficient condition for stability is $\lambda_n < \mu_n$, for each CR U_n; and a necessary condition for stability is $\lambda_n \leq \mu_n$ (cf. Fig. 1).

To complete the formulation, consider N_R actual or potential PU receivers, whose locations have been estimated via sensing [24], and let $\iota_R^{\max}$ denote the maximum average interference that can be tolerated by PU receiver R [7], [25]. Further, let $\mathcal{N}_R \subseteq \{U_n\}_{n=1}^N$ be the (sub-)set of CR nodes located in the proximity of PU R (not necessarily the entire CR network, as some CRs may be sufficiently far apart and do not interfere with PU R). Transmissions by CR U_n undergo

random shadowing and small-scale fading effects before arriving at close-by PU nodes. Approximate the channel gain $g_{n\to R}$ between CR U_n and the PU R as log-normal distributed [17], and define a binary random variable $a_n \in \{0, 1\}$, independent of $g_{n\to R}$, taking the value 1 with probability μ_n, and 0 with probability $1 - \mu_n$. Then, the average interference experienced at PU R is given by ($\kappa := 0.1 \ln(10)$)

$$\iota_R := \mathbb{E}\left\{ \sum_{n\in\mathcal{N}_R} a_n p_n g_{n\to R} \right\}$$

$$= \sum_{n=1}^N \mu_n e^{\kappa P_n + \kappa m_{n\to R} + \frac{\kappa^2}{2}\sigma_{n\to R}^2} \leq \iota_R^{\max}. \quad (4)$$

Variables $\{P_n, \mu_n, \rho_n, \lambda_n\}$, and $\{t_{n\to i}\}$ satisfying the constraints (3) and (4) can be supported by the wireless CR network. It is certainly desirable to design the network by selecting a feasible set of variables that are optimal in some sense. To this end, consider a concave utility $\mathcal{U}_n(\rho_n)$, and a convex cost $\mathcal{C}_n(P_n)$, representing the reward of exogenous rate ρ_n and the cost of power P_n for node U_n, respectively [26]. Notice that ρ_n is the average rate of packets generated at the application layer of node U_n to be eventually delivered to the sink U_{N+1} [22], [26]; thus, ρ_n represents an end-to-end performance metric. Capitalizing on the statistical description of SINRs and CR-to-PU channels, the *statistical* routing problem is formulated as:

$$(\text{P1}) \max_{\substack{\{P_n\},\{\rho_n\geq 0\}, \\ \{\mu_n\geq 0\},\{\lambda_n\geq 0\} \\ \{t_{n\to i}\geq 0\}}} \sum_{n=1}^N \mathcal{U}_n(\rho_n) - \sum_{n=1}^N \mathcal{C}_n(P_n) \quad (5a)$$

subject to

$$\rho_n + \sum_{i\in\mathcal{N}_{\to n}} \lambda_i t_{i\to n} r_{i\to n} \leq \lambda_n \sum_{j\in\mathcal{N}_{n\to}} t_{n\to j} r_{n\to j}$$
$$\forall n = 1, \ldots, N \quad (5b)$$

$$\sum_{i\in\mathcal{N}_{n\to}} t_{n\to i} \leq 1, \qquad \forall n = 1, \ldots, N \quad (5c)$$

$$\lambda_n \leq \mu_n, \ \mu_n \leq 1, \ \forall n = 1, \ldots, N \quad (5d)$$

$$P_n \leq P_n^{\max}, \qquad \forall n = 1, \ldots, N \quad (5e)$$

$$\iota_R^{\max} \geq \sum_{n\in\mathcal{N}_R} \mu_n e^{\kappa P_n + \kappa m_{n\to R} + \frac{\kappa^2}{2}\sigma_{n\to R}^2},$$
$$\forall R = 1, \ldots, N_R \quad (5f)$$

with $\{r_{n\to j}\}$ given by (2), and $P_n^{\max} := 10 \log_{10} p_n^{\max}$.

The non-convexity of constraints (5b) and (5f) makes problem (P1) non-convex, and thus hard to solve. Furthermore, function $Q(\cdot)$ in (2) is difficult to handle in an optimization problem. In the next section, an approximate but efficiently solvable version of (P1) will be formulated. But first, some remarks are in order.

Remark 1 (*Monotonically non-decreasing utilities*). It follows from [27, Thm. 5], that (P1) is optimally solved by setting $\{\lambda_n = \mu_n\}$ if each utility function $\mathcal{U}_n(\rho_n)$ is monotonically non-decreasing. As many practical utilities satisfy this condition, $\mathcal{U}_n(\rho_n)$ will be hereafter assumed non-decreasing, and variables $\{\lambda_n\}$ will be dropped. Strictly speaking, the choice

[1]"Exogenous" packets of a CR node are those generated from its application layer. On the other hand, "endogenous" packets refer to those received from the neighboring nodes of a CR node, and are to be routed by the network layer; see also Fig. 1.

$\{\lambda_n = \mu_n\}$ will lead to a solution of (P1) where queues may or may not be stable [23]. On the other hand, condition $\lambda_n < \mu_n$ is challenging because it entails an open constraint set. From a practical perspective, queue stability can be readily ensured by imposing in (5e) the condition $\lambda_n + \epsilon \leq \mu_n$, with $0 < \epsilon \ll 1$ small enough, and replacing variables $\{\mu_n\}$ with $\{\bar{\mu}_n := \mu_n - \epsilon\}$. □

Remark 2 *(Conventional multi-hop networks).* The proposed routing framework can be considered also for non-CR multihop random access networks when node-to-node channels can not be estimated accurately - what could emerge with e.g., a mobile ad hoc topology. Optimal routes and link reliabilities can be obtained by solving (P1), after discarding the interference constraints (5f), and re-defining the signal-to-noise ratio (SNR) of link $U_n \to U_i$ as $\gamma_{n \to i} = p_n g_{n \to i} / \sigma_i^2$. □

Remark 3 *(MAC protocol).* Since a random access protocol is adopted, a packet is as usual deemed lost when collisions among CR transmissions occur [cf. (2)], and no mutual interference is explicitly modeled in (1). However, the solution approach presented in the ensuing section can be effectively employed when different MAC strategies such as, e.g., carrier sensing medium access and orthogonal transmissions are utilized by the CR nodes; see also [26]. □

III. TRACTABLE ROUTING PROTOCOL

To convexify constraint (5f) it suffices to consider the logarithmic change of variables $\tilde{\mu}_n := \ln(\mu_n)$. As for the flow constraint (5b), consider first introducing auxiliary variables $\{\nu_n\}$ representing the probability of CRs to remain silent, together with the extra constraints $\mu_n + \nu_n = 1$, for $n = 1, \ldots, N$. Further, a simple way to obtain a tractable approximation of $Q(x)$ consists in exploiting the commonly used upper and lower bounds proposed in [28], [29], which are very tight for $x > \sqrt{2}/2$. Taking advantage of these bounds, and performing again a logarithmic change of variables $\tilde{\nu}_n := \ln(\nu_n)$, the probability $r_{n \to i}$ can be (tightly) bounded as

$$r_{n \to i} \geq e^{\sum_{j \in \mathcal{I}_{ni}} \tilde{\nu}_i} \times \left(1 - \frac{1}{12}e^{-\frac{1}{2}\left(\frac{P_n + m_{n \to i} - \bar{\Gamma}_{n \to i}}{\sigma_{n \to i}}\right)^2}\right.$$
$$\left. - \frac{1}{4}e^{-\frac{2}{3}\left(\frac{P_n + m_{n \to i} - \bar{\Gamma}_{n \to i}}{\sigma_{n \to i}}\right)^2}\right) \quad (6)$$

$$r_{n \to i} \leq e^{\sum_{j \in \mathcal{I}_{ni}} \tilde{\nu}_i} \times \left(1 - \alpha_1 e^{-\alpha_2\left(\frac{P_n + m_{n \to i} - \bar{\Gamma}_{n \to i}}{\sigma_{n \to i}}\right)^2}\right) \quad (7)$$

where $\alpha_1 = 0.28$, and $\alpha_2 = 0.64$ [29]. The premise for adopting the aforesaid bounds is that the decoding rate of CR links is at least ≈ 0.7. This condition is met in practice if CRs and PUs are sufficiently far apart (see, e.g. [19]). Furthermore, maximum packet error rates required for data and speech transmissions are generally considerably lower than 0.3 [6].

Consider now using the upper bound (7) for the incoming traffic, and the lower bound (6) for the outgoing flows. As (6) and (7) are tight, this replacement not only yields a tractable optimization problem, but also does not sacrifice optimality of the outcoming rates. With the logarithmic change of variable $\tilde{t}_{n \to i} = \ln(t_{n \to i})$, and after introducing auxiliary variables

$\{\breve{y}_{n \to i} \geq \sqrt{2}/2\}$ and $\{\hat{y}_{n \to i} \geq \sqrt{2}/2\}$, constraint (5b) can be approximated as

$$\rho_n + \sum_{i \in \mathcal{N}_{n \to}} e^{\tilde{\mu}_n + \tilde{t}_{n \to i} + \sum_{m \in \mathcal{I}_{ni}} \tilde{\nu}_m} \left(\frac{1}{12}e^{-\frac{1}{2}\hat{y}_{n \to i}} + \frac{1}{4}e^{-\frac{2}{3}\hat{y}_{n \to i}}\right)$$
$$+ \sum_{j \in \mathcal{N}_{\to n}} e^{\tilde{\mu}_j + \tilde{t}_{j \to n} + \sum_{m \in \mathcal{I}_{jn}} \tilde{\nu}_m} - \sum_{i \in \mathcal{N}_{n \to}} e^{\tilde{\mu}_n + \tilde{t}_{n \to i} + \sum_{m \in \mathcal{I}_{ni}} \tilde{\nu}_m}$$
$$- \alpha_1 \sum_{j \in \mathcal{N}_{\to n}} e^{\tilde{\mu}_j + \tilde{t}_{j \to n} + \sum_{m \in \mathcal{I}_{jn}} \tilde{\nu}_m - \alpha_2 \breve{y}_{j \to n}} \leq 0 \quad (8)$$

with the auxiliary constraints

$$\sigma_{n \to i}^{\frac{1}{2}}(\hat{y}_{n \to i}) \leq P_n + m_{n \to i} - \bar{\Gamma}_{n \to i} \quad (9)$$
$$\sigma_{i \to n}^{\frac{1}{2}}(\breve{y}_{i \to n}) \geq P_i + m_{i \to n} - \bar{\Gamma}_{i \to n}. \quad (10)$$

For notational convenience, define the variable vector $\mathbf{x}_n := [P_n, \{\tilde{t}_{n \to i}\}, \tilde{\mu}_n, \tilde{\nu}_n, \{\breve{y}_{j \to n}, \hat{y}_{n \to i}\}]^T$ per node U_n, for $n = 1, \ldots, N$. Upon re-expressing in a compact form the flow constraint (8) as $f_n(\{\mathbf{x}_n\}) \leq 0$, and defining the constraint set $\mathcal{B}_{\mathbf{x}_n}$ per node U_n as

$$\mathcal{B}_{\mathbf{x}_n} := \left\{\mathbf{x}_n : \sum_{i \in \mathcal{N}_{n \to}} e^{\tilde{t}_{n \to i}} \leq 1, e^{\tilde{\mu}_n} + e^{\tilde{\nu}_n} \leq 1,\right.$$
$$\left. P_n \leq P_n^{\max}, \text{and (9), (10)}\right\} \quad (11)$$

where (non)negativity of the variables is left implicit, problem (P1) can be re-formulated as

$$\text{(P2)} \quad \max_{\{\mathbf{x}_n \in \mathcal{B}_{\mathbf{x}_n}\}} \sum_{n=1}^{N} \mathcal{U}_n(\rho_n) - \sum_{n=1}^{N} \mathcal{C}_n(P_n) \quad (12a)$$

subject to $f_n(\{\mathbf{x}_n\}) \leq 0, \quad n = 1, \ldots, N \quad (12b)$

$$\sum_{n \in \mathcal{N}_R} f_{\iota_R,n}(\mathbf{x}_n) \leq \iota_R^{\max}, R = 1, \ldots, N_R \quad (12c)$$

where $f_{\iota_R,n}(\mathbf{x}_n) := e^{\tilde{\mu}_n + \kappa P_n + \kappa m_{n \to R} + \frac{\kappa^2}{2}\sigma_{n \to R}^2}$.

Constraints (8) are still non-convex because the last two sums (with their signs) are concave, and likewise (9) is also concave. Nevertheless, the structure of (P2) allows convex approximation methods for obtaining its solution efficiently. Among candidate methods, the successive convex approximation approach [30] is well suited for the problem at hand because it guarantees first-order KKT optimality under mild regularity conditions.

A. KKT solution via successive convex approximation

The general successive convex approximation method is outlined first. Suppose that the objective function to be maximized is concave in the optimization variables $\mathbf{x}$, and the constraint set is the intersection of a set $\mathcal{A} := \{\mathbf{x}|f_n(\mathbf{x}) \leq 0, n = 1, \ldots, N\}$ with a convex set $\mathcal{B}$, which captures convex constraints, if any. Assume that $f_n(\mathbf{x})$, $n = 1, \ldots, N$, are differentiable but generally non-convex functions. Then, starting from a feasible point $\mathbf{x}^{(0)} \in \mathcal{A} \cap \mathcal{B}$, a series $\ell = 1, \ldots$, of surrogate problems is solved, where $\mathcal{A}$ is substituted per iteration ℓ by a convex set $\mathcal{A}^{(\ell)}$. Since the intersection of convex sets yields a convex set, the resulting optimization problems are convex. For each $n = 1, \ldots, N$, let $\tilde{f}_n(\mathbf{x}; \mathbf{x}^{(\ell)})$ denote the surrogate convex function for $f_n(\mathbf{x})$, which may

depend on the solution $\mathbf{x}^{(\ell)}$ to the problem of the previous $(\ell-1)$-st iteration. Then, the convex set $\mathcal{A}^{(\ell)}$ is constructed as $\mathcal{A}^{(\ell)} := \{\mathbf{x}|\tilde{f}_n(\mathbf{x};\mathbf{x}^{(\ell)}) \leq 0, n = 1, \ldots, N\}$. Provided that each function $\tilde{f}_n(\mathbf{x};\mathbf{x}^{(\ell)})$, $n = 1, \ldots, N$, is convex, differentiable, and satisfies conditions [30]

c1) $f_n(\mathbf{x}) \leq \tilde{f}_n(\mathbf{x};\mathbf{x}^{(\ell)})$, $\quad \forall \mathbf{x} \in \mathcal{A}^{(\ell)} \cap \mathcal{B}$
c2) $f_n(\mathbf{x}^{(\ell)}) = \tilde{f}_n(\mathbf{x}^{(\ell)};\mathbf{x}^{(\ell)})$, and
c3) $\nabla f_n(\mathbf{x}^{(\ell)}) = \nabla \tilde{f}_n(\mathbf{x}^{(\ell)};\mathbf{x}^{(\ell)})$

the series of solutions to the approximate problems converge to the KKT point of (P2) [30].

In order to apply the successive convex approximation method to (P2), surrogate constraints for the non-convex constraints must be determined. The first three terms in (8) are convex, whereas the fourth and fifth terms are concave. Letting $-e^{x_1+\beta x_2-\alpha x_3}$ represent one of the non-convex summands, a convex surrogate function satisfying c1)-c3) can be obtained by replacing the non-convex summands with the affine function

$$- e^{x_1+\beta x_2-\alpha x_3} \leq e^{x_1^{(\ell)}+\beta x_2^{(\ell)}-\alpha x_3^{(\ell)}} \times \left[(x_1^{(\ell)} - x_1)\right.$$
$$\left. + \beta(x_2^{(\ell)} - x_2) - \alpha(x_3^{(\ell)} - x_3) - 1\right]. \quad (13)$$

As for (9), an upper-bound of $\sqrt{\hat{y}_{n\to j}}$ can be obtained via the supporting hyperplane, and the resulting surrogate convex constraints become

$$\frac{\hat{y}_{n\to j} - \hat{y}_{n\to j}^{(\ell)}}{2\sqrt{\hat{y}_{n\to j}^{(\ell)}}} + \sqrt{\hat{y}_{n\to j}^{(\ell)}} - P_n - m_{n\to i} + \bar{\Gamma}_{n\to i} \leq 0. \quad (14)$$

Overall, the problem to solve in the ℓ-th iteration is given by (P2) with (9) replaced by (14) to form the surrogate constraint set $\tilde{\mathcal{B}}_n$, and by employing (13) along the feasible points $\{\mathbf{x}_n^{(\ell)}\}_{n=1}^N$ to obtain a surrogate convex flow conservation constraint $\tilde{f}_n^{(\ell)}(\{\mathbf{x}_n\}) \leq 0$; that is,

$$(\text{P2}^{(\ell)}) \quad \max_{\{\mathbf{x}_n \in \tilde{\mathcal{B}}_n\}} \sum_{n=1}^N \mathcal{U}_n(\rho_n) - \sum_{n=1}^N \mathcal{C}_n(P_n) \quad (15a)$$
$$\text{subject to (12c) and } \tilde{f}_n^{(\ell)}(\{\mathbf{x}_n\}) \leq 0, \forall n \quad (15b)$$

Problem $(\text{P2}^{(\ell)})$ is convex, and thus efficiently solvable using interior-point methods [31]. It is worth mentioning that the solution of $(\text{P2}^{(\ell)})$, $\ell = 1, 2, \ldots$ always lies inside the feasibility region of the original non-convex problem (P2) [30]. This observation suggests readily an *online* practical implementation of the algorithm whereby node U_n does not necessarily wait for the successive convex approximation algorithm to converge, but rather relies on $\mathbf{x}_n^{(\ell)}$ as and when it becomes available. In the limit (i.e., for $\ell \gg 1$), $\mathbf{x}_n^{(\ell)}$ will be KKT-optimal. An online implementation of the iterative optimization allows tracking of slow variations in the network topology and SINR statistics.

IV. DISTRIBUTED STATISTICAL ROUTING

To obviate the high communication cost associated with the collection of channel statistics for all links at a central processing unit, and the subsequent dissemination of the optimized variables, it is of prime interest to solve (P2) in a distributed manner. A distributed cross-layer optimization algorithm is also desirable because of its scalability with regards to power requirements and network size, and robustness to isolated points of failure.

Distributing (P2) is tantamount to developing a distributed solver for each of the convex problems $(\text{P2}^{(\ell)})$, $\ell = 1, 2, \ldots$. To this end, it is necessary to decompose $(\text{P2}^{(\ell)})$ into smaller sub-problems, which can be locally solved by nodes $\{U_n\}$ via *local* message exchanges. Unfortunately, the interference constraints (12c) challenge decomposability, as they couple portions of the CR network. Furthermore, for each U_n, constraint (15b) involves variables pertaining to the one-hop neighboring nodes $U_i \in \mathcal{N}_{\to n}$, and to CRs in the collision-related sets $\{\mathcal{I}_{jn}\}$ and $\{\mathcal{I}_{ni}\}$. To overcome the first hurdle, consider first the following problem

$$(\text{P3}^{(\ell)}) \quad g(\{\iota_{R,n}^{\max}(\ell)\}) :=$$
$$\max_{\{\mathbf{x}_n \in \tilde{\mathcal{B}}_n\}} \sum_{n=1}^N \mathcal{U}_n(\rho_n) - \sum_{n=1}^N \mathcal{C}_n(P_n) \quad (16a)$$
$$\text{subject to } \tilde{f}_n(\{\mathbf{x}_n\}) \leq 0, \quad \forall n \quad (16b)$$
$$f_{\iota_R,n}(\mathbf{x}_n) \leq \iota_{R,n}^{\max}(\ell), \ \forall n, R \quad (16c)$$

where the interference $\iota_R^{\max}$ for PU R is pre-partitioned in *given* per-CR fractions $\{\{\iota_{R,n}^{\max}(\ell)\}\}_{n\in\mathcal{N}_R}$. Problem $(\text{P2}^{(\ell)})$ will be revisited later on. Then, collect *local* copies of $\mathbf{x}_{j\to n} := [\tilde{t}_{j\to n}, P_j, \tilde{\mu}_j]^T$ at node U_n, for each $j \in \mathcal{N}_{\to n}$ into a vector $\mathbf{x}_{j\to n,n} := [\tilde{t}_{j\to n,n}, P_{j,n}, \tilde{\mu}_{j,n}]^T$. Likewise, let $\{\tilde{\nu}_{n,m}\}$ denote local copies of $\{\tilde{\nu}_m|m \in \mathcal{I}_n\}$, with $\mathcal{I}_n := (\cup_{i\in\mathcal{N}_{n\to}}\mathcal{I}_{ni})\cup(\cup_{j\in\mathcal{N}_{\to n}}\mathcal{I}_{jn})$; i.e., local copies of $\tilde{\nu}_m$ for users that may interfere with U_n's transmissions. Then, $(\text{P3}^{(\ell)})$ can be equivalently re-formulated as

$$(\text{P4}^{(\ell)}) \quad \max_{\{\mathbf{x}_n \in \tilde{\mathcal{B}}_{\mathbf{x}_n}\}} \sum_{n=1}^N \mathcal{U}_n(\rho_n) - \sum_{n=1}^N \mathcal{C}_n(P_n) \quad (17a)$$
subject to
$$\tilde{f}_n(\mathbf{x}_n, \{\mathbf{x}_{j\to n,n}\}, \{\nu_{m,n}\}) \leq \ \forall n \quad (17b)$$
$$f_{\iota_R,n}(\mathbf{x}_n) \leq \iota_{R,n}^{\max}(\ell), \quad \forall n, R \quad (17c)$$
$$\mathbf{x}_{j\to n} = \mathbf{x}_{j\to n,n}, \quad j \in \mathcal{N}_{\to n}, \forall n, \quad (17d)$$
$$\tilde{\nu}_m = \tilde{\nu}_{m,n}, \quad m \in \mathcal{I}_n, \forall n \quad (17e)$$

where the notation $\tilde{f}_n(\mathbf{x}_n, \{\mathbf{x}_{j\to n,n}\}, \{\nu_{m,n}\})$ emphasizes the dependence of the surrogate flow conservation constraint $\tilde{f}_n(\cdot)$ on the newly introduced local variables. Problem $(\text{P4}^{(\ell)})$ is amenable to a distributed solution, where (17d)-(17e) can be enforced by means of local message passing.

Suppose that there is a non-zero probability (possibly multi-hop) directed path connecting U_n to nodes $U_m \in \mathcal{N}_{\to n} \cup \mathcal{N}_{n\to} \cup \mathcal{I}_n$; i.e., nodes coupled in the optimization problem. If not, a control channel can be employed as usual. Problem $(\text{P4}^{(\ell)})$ may be solved in a distributed manner using the dual sub-gradient method [32], [33]. However, recovering the primal variables $\{\mathbf{x}_n\}$ from the Lagrange multipliers optimizing the dual function is not always guaranteed if the objective in (17a) is not strictly convex, and the step-size in the sub-gradient ascent is constant. Furthermore, primal averaging can not be performed in this case, unless the equality constraints are appropriately relaxed [34].

Algorithm 1 Distributed algorithm for (P4$^{(\ell)}$)

Assumption: bidirectional links, or bidirectional control channels.

Use solution of (P4$^{(\ell-1)}$) to initialize variables.

for $l = 0, 1, \ldots$ (repeat until convergence) **do**

  Receive multipliers $\{\mathbf{q}_{n,i}(l)\}_{i \in \mathcal{N}_{n \to}}, \{v_{n,p}(l)\}_{p \in \{p | U_n \in \mathcal{I}_p\}}$.

  Update $\bar{\mathbf{x}}_n := \{\mathbf{x}_n, \{\mathbf{x}_{j \to n,n}\}, \{\nu_{m,n}\}\}$ via (19a)

  Transmit $\mathbf{x}_{n \to i}(l)$ to $U_i \in \mathcal{N}_{n \to}$, and $\mathbf{x}_{j \to n,n}(l)$ to $U_j \in \mathcal{N}_{\to n}$

  Transmit $\{\tilde{\nu}_{m,n}\}$ to $U_m \in \mathcal{I}_n$ via neighboring nodes. Forward $\{\tilde{\nu}_{m,j}\}_{j \in \mathcal{N}_{\to n}}$ to $U_i \in \mathcal{N}_{n \to}$.

  Receive $\mathbf{x}_{j \to n}(l)$ from $U_i \in \mathcal{N}_{n \to}$, and $\mathbf{x}_{n \to i,i}(l)$ from $U_i \in \mathcal{N}_{n \to}$

  Receive $\tilde{\nu}_{n,p}(l)$ from $U_p \in \{U_p | U_n \in \mathcal{I}_p\}$

  Dual update via (19c)-(19d).

  Transmit multipliers $\mathbf{q}_{n,i}(l+1)$ to $U_i \in \mathcal{N}_{n \to}$.

  Transmit $v_{m,n}(l+1)$ to $U_m \in \mathcal{I}_n$ via neighboring nodes, forward $\{v_{m,j}\}_{j \in \mathcal{N}_{\to n}}(l+1)$ to $U_i \in \mathcal{N}_{n \to}$.

  Use parameters $\bar{\mathbf{x}}_n$ to transmit data in case of on-line implementation.

end for

One effective remedy is offered by the alternating direction method of multipliers (ADMoM), where the optimization argument in (P4$^{(\ell)}$) is augmented with a quadratic regularization term corresponding to the squared norm of the equality constraints [35, Sec. 3.4]. Specifically, letting $\{\mathbf{q}_{j,n}\}$ and $\{v_{m,n}\}$ denote the multipliers associated with the equality constraints (17d) and (17e), respectively, the partial quadratically-augmented Lagrangian function is given by

$$
\mathcal{L}(\{\bar{\mathbf{x}}_n\}, \{\mathbf{q}_{j,n}\}, \{v_{m,n}\}) := \sum_{n=1}^{N} \Big[-\mathcal{U}_n(\rho_n) + \mathcal{C}_n(P_n)
$$
$$
+ \sum_{j \in \mathcal{N}_{\to n}} \Big(\mathbf{q}_{j,n}^T (\mathbf{x}_{j \to n} - \mathbf{x}_{j \to n,n}) + \frac{c}{2} \|\mathbf{x}_{j \to n} - \mathbf{x}_{j \to n,n}\|_2^2 \Big)
$$
$$
+ \sum_{m \in \mathcal{I}_n} \Big(v_{m,n}(\tilde{\nu}_m - \tilde{\nu}_{m,n}) + \frac{c}{2}(\tilde{\nu}_m - \tilde{\nu}_{m,n})^2 \Big) \Big] \quad (18)
$$

where $\bar{\mathbf{x}}_n := \{\mathbf{x}_n, \{\mathbf{x}_{j \to n,n}\}, \{\nu_{m,n}\}\}$, and $c > 0$ is an arbitrary constant. Notice that $\mathcal{L}(\cdot)$ is defined over the primal feasible region $\mathcal{A} := \cap_{n=1}^{N} \mathcal{A}_n$, with $\mathcal{A}_n := \{\bar{\mathbf{x}}_n | \mathbf{x}_n \in \tilde{\mathcal{B}}_{\mathbf{x}_r}, (17b), (17c)\}$. ADMoM amounts to performing the following iterations (l denotes the iteration index)

[I.1] *Primal update.* Given $\{\mathbf{q}_{j,n}(l)\}$ and $\{v_{m,n}(l)\}$, update primal variables in a coordinate descent fashion; i.e., for $n = 1, \ldots, N$, update $\bar{\mathbf{x}}_n$ as:

$$
\bar{\mathbf{x}}_n(l+1) := \min_{\bar{\mathbf{x}}_n \in \mathcal{A}_n} \mathcal{L}_n(\bar{\mathbf{x}}_n, l) \quad (19a)
$$
$$
\mathcal{L}_n(\bar{\mathbf{x}}_n, l) := \mathcal{L}(\bar{\mathbf{x}}_1(l), \ldots, \bar{\mathbf{x}}_{n-1}(l), \bar{\mathbf{x}}_n, \bar{\mathbf{x}}_{n+1}(l), \ldots, \bar{\mathbf{x}}_N(l),
$$
$$
\{\mathbf{q}_{j,n}(l)\}, \{v_{m,n}(l)\}) \quad (19b)
$$

where $\mathcal{L}_n(\bar{\mathbf{x}}_n, l)$ is obtained by keeping $\{\bar{\mathbf{x}}_m(l)\}_{m \neq n}$ fixed to their values at iteration l.

[I.2] *Dual update.* Given the primal variables $\{\bar{\mathbf{x}}_n(l+1)\}_{n=1}^{N}$, updated multipliers as:

$$
\mathbf{q}_{j,n}(l+1) = \mathbf{q}_{j,n}(l) + \beta[\mathbf{x}_{j \to n}(l+1) - \mathbf{x}_{j \to n,n}(l+1)] \quad (19c)
$$
$$
v_{m,n}(l+1) = v_{m,n}(l) + \beta[\tilde{\nu}_m(l+1) - \tilde{\nu}_{m,n}(l+1)] \quad (19d)
$$

where $\beta > 0$ is the step-size.

Once the primal iterates of the neighboring nodes $\{\mathbf{x}_{j \to n}(l+1)\}$ and $\{\tilde{\nu}_m(l+1)\}$ become available at node U_n, the dual updates (19c)-(19d) can be performed locally. As for the primal update, the local augmented Lagrangian [cf. (18)]

$$
\mathcal{L}_n(\bar{\mathbf{x}}_n, l) = -\mathcal{U}_n(\rho_n) + \mathcal{C}_n(P_n)
$$
$$
+ \sum_{j \in \mathcal{N}_{\to n}} \Big[-\mathbf{q}_{j,n}^T(l)\mathbf{x}_{j \to n,n} + \frac{c}{2}\|\mathbf{x}_{j \to n}(l) - \mathbf{x}_{j \to n,n}\|_2^2 \Big]
$$
$$
+ \sum_{i \in \mathcal{N}_{n \to}} \Big[\mathbf{q}_{n,i}^T(l)\mathbf{x}_{n \to i} + \frac{c}{2}\|\mathbf{x}_{n \to i} - \mathbf{x}_{n \to i,i}(l)\|_2^2 \Big]
$$
$$
+ \sum_{m \in \mathcal{I}_n} \Big[-v_{m,n}(l)\tilde{\nu}_{m,n} + \frac{c}{2}(\tilde{\nu}_m(l) - \tilde{\nu}_{m,n})^2 \Big]
$$
$$
+ \sum_{p|n \in \mathcal{I}_p} \Big[v_{n,p}(l)\tilde{\nu}_n + \frac{c}{2}(\tilde{\nu}_n - \tilde{\nu}_{n,p}(l))^2 \Big] \quad (20)
$$

can be minimized at node U_n upon collecting $\mathbf{x}_{n \to i,i}(l)$ and multipliers $\{\mathbf{q}_{n,i}(l)\}$ from the one-hop neighboring nodes $U_i \in \mathcal{N}_{n \to}$, and $\{v_{n,p}(l)\}$ and $\tilde{\nu}_{n,p}(l)$ from nodes $U_p \in \{U_p | U_n \in \mathcal{I}_p\}$; that is, from the nodes whose transmissions can collide with the ones of U_n. Roughly speaking, the latter quantities pertain to the two-hop neighborhood of node U_n and are due to the basic properties of the random access strategy. If a different medium access protocol such as, e.g., CSMA is employed, (17e) will not be required and the message-passing overhead can be further reduced. The ADMoM-based distributed algorithm is tabulated as in Algorithm 1, and the convergence to the optimal primal arguments $\{\bar{\mathbf{x}}_n(l)\}$ as $l \to \infty$ is summarized in the following proposition.

Proposition 1. *If there exists a non-zero probability (possibly multi-hop) directed path connecting U_n to nodes $U_m \in \mathcal{N}_{\to n} \cup \mathcal{N}_{n \to} \cup \mathcal{I}_n$, for all n, the iterates $\{\bar{\mathbf{x}}_n(l)\}$ generated by Algorithm 1 converge to a globally optimal solution to (P4$^{(\ell)}$).*

Proof. Existence of a path connecting U_n to nodes $U_m \in \mathcal{N}_{\to n} \cup \mathcal{N}_{n \to} \cup \mathcal{I}_n$, guarantees a regular exchange of local primal variables and multipliers among neighboring nodes. Under this assumption, convergence of the primal iterates $\{\bar{\mathbf{x}}_n(l)\}$ to their optimal values as $l \to \infty$ can be readily established using the result in [35, Prop. 4.2]. □

Algorithm 1 can also be implemented in an online fashion. The equality constraint violation during the initial iterations of the algorithm may induce an initial increase of some queues. Thus, an online implementation is feasible if nodes can afford such a potential increase in the queue length before reaching consensus on the local variables.

A. Handling the interference constraint via primal decomposition

Reconsider now problem (P2$^{(\ell)}$), where the interference budgets $\{\iota_R^{\max}\}_{R=1}^{N_R}$ are *not* partitioned *a priori* among CR nodes. As primal variables become feasible only when dual decomposition algorithms have converged, utilization of network parameters obtained from intermediate iterates can possibly lead to violation of the interference constraint. To enforce strict PU protection during network operation, the primal decomposition technique is invoked here; see, e.g., [33]. With this method, resources shared among CR nodes are essentially allocated by a master problem. Specifically, at each

iteration $k = 1, 2, \ldots$ of the primal decomposition algorithm, problem $(P3^{(\ell,k)})$ is solved for given $\{\{\iota_{R,n}^{\max}(\ell,k)\}_{R,n}\}$; then, $\{\{\iota_{R,n}^{\max}(\ell,k)\}_{R,n}\}$ are updated by solving the following problem:

$$(P5^{(\ell,k)}) \qquad \{\iota_{R,n}^{\max}(\ell,k)\} = \arg \max_{\{\iota_{R,n}^{\max}\}} g(\{\iota_{R,n}^{\max}\}) \qquad (21a)$$

$$\text{subject to } \iota_{R,n}^{\max} \geq 0, \qquad (21b)$$

$$\sum_{n \in \mathcal{N}_R} \iota_{R,n}^{\max} \leq \iota_R^{\max}, \, \forall \, R. \qquad (21c)$$

To solve $(P5^{(\ell,k)})$, the subgradient algorithm can be employed [31]. Specifically, the subgradient of $g(\{\iota_{R,n}^{\max}(\ell,k)\})$ with respect to $\iota_{R,n}^{\max}(\ell,k)$ is given by the *optimal* Lagrange multiplier $u_{R,n}(k)$ corresponding to the constraint $f_{\iota_R,n}(\mathbf{x}_n) \leq \iota_{R,n}^{\max}(\ell,k)$ in $(P3^{(\ell)})$ at iteration k [33]. Therefore, $\iota_{R,n}^{\max}(\ell,k)$ is updated as

$$\iota_{R,n}^{\max}(\ell,k+1) = \mathcal{P}_{\iota_R}\left\{\iota_{R,n}^{\max}(\ell,k) + \xi(k+1)u_{R,n}(k)\right\} \quad (22)$$

where $\xi(\cdot)$ is the step size, and $\mathcal{P}_{\iota_R}\{\cdot\}$ denotes projection onto the region defined by (21c), operation that can be efficiently computed as in, e.g. [36]. At each step k of the primal algorithm, CR nodes can employ variables obtained from $(P3^{(\ell,k)})$ for network operation, as PU interference protection is enforced by updates (22).

The projection in (22) needs to be performed by a "head node" in the CR sub-network CR $\mathcal{N}_R$, which is formed by nodes that are coupled by constraint $\sum_{n \in \mathcal{N}_R} f_{\iota_R,n}(\mathbf{x}_n) \leq \iota_{R,n}^{\max}(\ell)$. Per iteration k, the head node has to collect the optimal Lagrange multipliers from the CRs in $\mathcal{N}_R$, and then broadcast the updated interference budgets $\{\iota_{R,n}^{\max}(\ell,k+1)\}_n$. This leads to a semi-distributed algorithm, but the high message-passing overhead entailed by centralized solutions is nonetheless alleviated. The online algorithm obtained through the successive convex approximation and primal decomposition is tabulated in Algorithm 2, and its convergence properties are summarized next.

Proposition 2. *If there is a cycle connecting nodes $U_n \in \mathcal{N}_R$, $\forall \, R$, the iterates generated by Algorithm 2 converge to a KKT solution to* $(P2)$.

Proof. Since the original problem $(P2^{(\ell,k)})$ is convex, the subproblems $(P3^{(\ell,k)})$ as well as the master problem $(P5^{(\ell,k)})$ are all convex, and thus the globally optimal solution of $(P2^{(\ell)})$ is attained via primal decomposition [33]. Existence of a cycle connecting nodes $U_n \in \mathcal{N}_R$ ensures that the multipliers $\{\xi(k)\}$ can be collected to a cluster head node, and that $\{\iota_{R,n}^{\max}(\ell,k+1)\}$ can be subsequently sent back. Finally, since $(P2^{(\ell)})$ is optimally solved per iteration ℓ of the successive convex approximation, convergence of Algorithm 2 to a KKT point of (P2) is guaranteed [30]. $\square$

Remark 4 *(Fully distributed algorithm).* At the expense of possibly sacrificing optimality of the resultant exogenous rates, coefficients $\{\iota_{R,n}^{\max}\}$ can be set a priori based on the distance between CRs and PU R. This may be reasonable especially if shadowing can not be estimated [24]. In this case, it is not necessary to compute the primal decomposition iterates. $\square$

Algorithm 2 Overall on-line algorithm for (P2)

Assumption: Path connecting all $U_n \in \mathcal{N}_R$, for all PUs $R = 1, \ldots, N_R$.
for $\ell = 1, \ldots$ (repeat until convergence) **do**
    Use solution of $(P2^{(\ell-1)})$ to compute (13)-(14). If $\ell = 1$, use suitable feasible point.
    for $k = 0, 1, \ldots$ (repeat until convergence) **do**
        Receive $u_{R,n}(k)$ from head node.
        Solve $(P3^{(\ell,k)})$ using Algorithm 1.
        Transmit multiplier $u_{R,n}(k)$ to head node via neighboring nodes.
        If head node: update $\{\iota_{R,n}^{\max}(\ell,k+1)\}$ via (22).
        Utilize $\bar{\mathbf{x}}_n(k)$ for network operation.
    end for
end for

V. PACKET DELIVERABILITY IN DYNAMIC CR ENVIRONMENTS

Statistics of the SINR may vary during network operation, because of the dynamic nature of shadow fading [24], and the variable PU interference levels [cf. (1)]. CR topology may also change with time. Proximity of PUs with intermittent activity, or, mobile PU devices may loose link connectivity during certain time intervals. The routing problem (P2) must be (re-)solved whenever network topology and SINR statistics change. Alternatively, it can be implemented online to track slow environmental dynamics. Either way, it is necessary to establish conditions ensuring that packets are eventually delivered to the sink when routes, MAC, and physical layer parameters are regularly updated.

Let $s_n(\tau) \in \{0,1\}$ be a binary variable taking value 1 if a packet, after having been randomly routed through the network, is placed in U_n's queue at time τ, and let $\vartheta_n(\tau) := \Pr\{s_n(\tau) = 1\}$ denote the probability of such an event. Further, collect $\{\vartheta_n(\tau)\}$ in the $(N+1) \times 1$ vector $\boldsymbol{\vartheta}(\tau) := [\vartheta_1(\tau), \ldots, \vartheta_{N+1}(\tau)]^T$. CR-PU hierarchy may prevent CR nodes from forwarding packets during certain time intervals. Let $\ell_{n \to j}$ be a binary variable that takes value 1 if link $U_n \to U_j$ is active, and define $\chi_{n \to j} := \Pr\{\ell_{n \to j} = 1\}$. If active, link $U_n \to U_j$ is characterized by a link reliability $r_{n \to j}(\tau)$. Probabilities $\{\chi_{n \to j}\}$ clearly depend on PU activity factors and locations, and determine the average connectivity of the CR network [11].

If a packet is in U_n's queue at time τ, then U_n may decide with probability $t_{n \to j}(\tau)$ to route it through one of the available links, where index τ emphasizes the time-variability of routes. Clearly, if neither node locations nor the PU interference or channel conditions change for a certain number of time slots, then $\{t_{n \to i}(\tau)\}$ and $\{r_{n \to i}(\tau)\}$ remain invariant. The evolution of $\{\vartheta_n(\tau)\}$ can thus be fully characterized by the product probabilities $\{t_{n \to j}(\tau)\, r_{n \to j}(\tau)\}$, and the link availability factors $\{\chi_{n \to j}\}$. Upon invoking the law of total probability, it holds that $\vartheta_n(\tau + 1) = \sum_{i=1}^{N+1} \Pr\{s_n(\tau + 1) = 1 | s_i(\tau) = 1, \ell_{i \to n} = 1\}\Pr\{s_i(\tau) = 1\}\Pr\{\ell_{i \to n} = 1\} = \sum_{i=1}^{N+1} t_{n \to j}(\tau)r_{n \to j}(\tau)\chi_{n \to j}\vartheta_i(\tau)$. Define the $(N+1) \times (N+1)$ packet delivery probability matrix $\mathbf{D}(\tau)$, whose off-diagonal entry (i, n) is $\{t_{n \to i}r_{n \to i}\chi_{n \to i}\}$ if U_i is a one-hop neighbor of U_n, and 0 otherwise. The diagonal entry (n, n) of $\mathbf{D}(\tau)$ represents the probability that a packet remains in

U_n's queue, which equals $1 - \sum_{i \neq n} t_{n \to i} r_{n \to i} \chi_{n \to i}$. Finally, since the sink node will not route packets to any other node, set the $(n, N+1)$-th entry of $\mathbf{D}(\tau)$ to $D_{n,N+1}(\tau) = 0$, and $D_{N+1,N+1}(\tau) = 1$. Matrix $\mathbf{D}(\tau)$ is by construction a column stochastic, meaning that $\mathbf{D}^T(\tau)\mathbf{1}_{N+1} = \mathbf{1}_{N+1}$ for all τ. Then, the evolution of $\{\vartheta_n(\tau)\}$ can be expressed in matrix-vector form as $\boldsymbol{\vartheta}(\tau + 1) = \mathbf{D}(\tau)\boldsymbol{\vartheta}(\tau)$.

Using an inductive argument, it is possible to show that the (i, n)th entry of the stochastic matrix $\bar{\mathbf{D}}(t) := \prod_{\tau=1}^{t} \mathbf{D}(\tau)$ represents the probability that a packet generated at U_n reaches node U_i in t time slots [37, Ch. 2]. Therefore, it readily follows that a packet is eventually delivered to the sink node U_{N+1} if and only if

$$\lim_{t \to +\infty} \boldsymbol{\vartheta}(t) = \lim_{t \to +\infty} \bar{\mathbf{D}}(t)\boldsymbol{\vartheta}(0) = [\mathbf{0}_N^T \; 1]^T \quad (23)$$

holds for any initial distribution $\boldsymbol{\vartheta}(0)$. A simple condition on the CR network topology is provided next in order for (23) to be satisfied.

Proposition 3. *If* $\sum_{i \neq n} \chi_{i \to n} > 0$, $\forall \{U_n\}_{n=1}^N$ *and* $\sum_n \chi_{n \to N+1} > 0$, *a packet stochastically routed according to probabilities* $\{t_{n \to i}(\tau)\}$ *over links with reliabilities* $\{r_{n \to i}(\tau)\}$ *will be eventually delivered to the destination with probability (w.p.)* 1.

Proof. The conditions of Prop. 3 ensure that there exists a multi-hop path connecting each node to the destination U_{N+1} in the average connectivity graph, where link $U_n \to U_j$ is present if $\chi_{n \to j} > 0$ [11]. Let t^* be the minimum number of time slots such that U_{N+1} can be reached from any node with non-zero probability; i.e., $t^* = \min\{t : \bar{D}_{N+1,n}(t) > 0 \; \forall n = 1, \ldots, N\}$. Then, the probability that a packet is in U_{N+1}'s queue at time $t^* + 1$ is given by

$$\vartheta_{N+1}(t^* + 1) = \sum_{n=1}^{N} \bar{D}_{N+1,n}(t^*)\vartheta_n(t^*) + \vartheta_{N+1}(t^*). \quad (24)$$

Arguing by contradiction, suppose that $\lim_{t^* \to +\infty} \vartheta_{N+1}(t^*) = \alpha < 1$; meaning that the packet is not delivered to U_{N+1} w.p. $1 - \alpha > 0$. Taking the limit on both sides of (24), one arrives at

$$\lim_{t^* \to +\infty} \vartheta_{N+1}(t^* + 1)$$

$$= \lim_{t^* \to +\infty} \left[\sum_{n=1}^{N} \bar{D}_{N+1,n}(t^*)\vartheta_n(t^*) + \vartheta_{N+1}(t^*) \right]$$

$$\geq \min_n \{\bar{D}_{N+1,n}(t^*)\} \sum_{n=1}^{N} \lim_{t^* \to +\infty} \vartheta_n(t^*) + \alpha. \quad (25)$$

But since $\sum_{n=1}^{N} \lim_{t^* \to +\infty} \vartheta_n(t^*) = 1 - \alpha > 0$ and $\min_n \{\bar{D}_{N+1,n}(t^*)\} > 0$, (25) can not hold, thus completing the proof. $\square$

Requiring the existence of an average node-to-destination multihop path is tantamount to having a Markov transition matrix with a unique absorbing state (the sink node) corresponding to the average graph [cf. (23)]. If a node U_n is able to receive packets, but cannot forward them to any other node due to a persistent activity of PU nodes in its proximity (which violates the condition of Proposition 3), then the constraints $t_{i \to n} = 0$ for all $i \in \mathcal{N}_{\to n}$ should be added in (P1).

VI. NUMERICAL RESULTS

Consider the scenario depicted in Fig. 2, where $N = 7$ CR nodes cooperate in routing packets to the destination U_8. Two PU sources also transmit to their intended receivers with power 10 dBW. In order to protect the PU system without knowing the locations of the PU receivers, 7 points on the boundary of the PUs' coverage regions are selected [24]. The PU interference threshold is set to -80 dBW. The path loss obeys the model $\|\mathbf{x}_n - \mathbf{x}_j\|^{-\eta}$, with $\eta = 3.5$. Log-normal shadowing is generated with standard deviation 6 dB, and $m = 1$ is used for the small-scale fading (Rayleigh) [16]. The maximum transmit-power of the CR nodes is set to $P_n^{\max} = 0$ dBW, and the noise power is 10^{-8} W. The SINR threshold $\bar{\Gamma}_n = -10$ dB, and the sum of exogenous rates $\sum_{n=1}^{N} \rho_n$ is maximized; that is $\mathcal{U}_n(\rho_n) = \rho_n$ and $\mathcal{C}_n(P_n) = 0$, for all $n = 1, \ldots, N$. A larger scale network could also be considered, but the conclusion that one could draw do not depend on the network size.

Fig. 2(a) depicts the optimal routing probabilities $\{t_{n \to i}\}$, obtained by solving (P2) with Algorithm 2. At the first iteration $\ell = 1$, a feasible starting point is obtained by properly modifying the approach of [38] to the problem at hand, and setting the step-size in (22) equal to 1. It can be seen that there is a tendency not to route packets through the "southern" region of the network; i.e., through nodes that are closer to the PU systems. For example, packets generated by U_2 are more likely to be routed through links $U_4 \to U_6$ and $U_6 \to U_7$, rather than choosing the shortest path $U_2 \to U_4 \to U_5 \to U_8$. Furthermore, node U_5 may decide to send packets to U_6 rather than attempting direct transmission to U_8 with considerably high probability. This is due to the fact that links starting from and ending to U_4 and U_5 are characterized by a higher fading- and interference-induced outage probability, as showed in Fig. 2(b). In fact, not only PU interference has a detrimental effect on the CR SINRs, but also U_2, U_4, and U_5 are confined to use a lower transmit-power in order to enforce protection of the PU receivers. Notice also that U_2 may decide to transmit to U_1 instead of U_4 with considerably high probability. On the other hand, packets generated by U_1 and U_3 are routed through U_7 with high probability, which in this case coincides also with the shortest path. Interestingly, it is necessary to use the primal decomposition algorithm only during the first 5-6 iterations out of the total 14 (on average) in the successive convex approximation algorithm. In fact, the per-CR interference levels quickly stabilize around steady-state values, with subtle variations for $\ell > 6$.

To verify adaptability of the routing probabilities and link reliabilities to the states of the PU systems, consider the case of Fig. 3(a), where the same CR network operates only with PU 1 present. Compared to Fig. 2(a), U_4 now forwards an increased amount of traffic through node U_5. As PU 2 is inactive, the outage probability of link $U_4 \to U_5$ is lower in this case, as confirmed by Fig. 3(b). Furthermore, U_5 can raise its transmit-power of 10 dB, which significantly decreases the outage probability of link $U_5 \to U_8$. As a result, almost none of the packets (2%) are sent to U_6. Finally, notice that CR U_6 now splits its traffic evenly between P_5 and P_7. The average exogenous traffic rates, averaged over 20 different
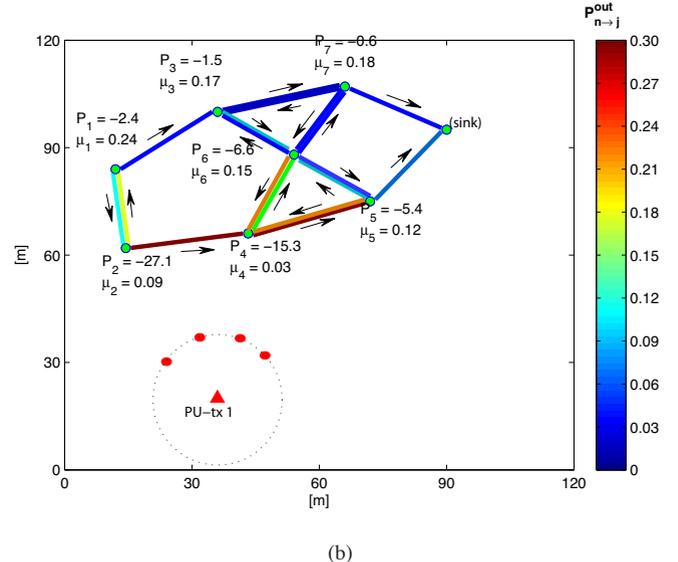
(a)
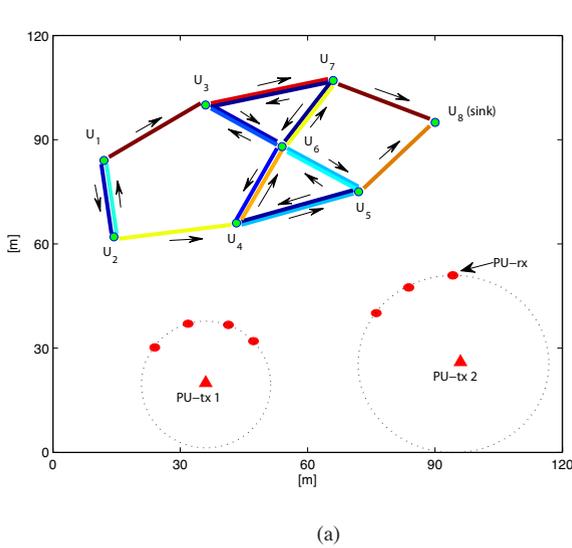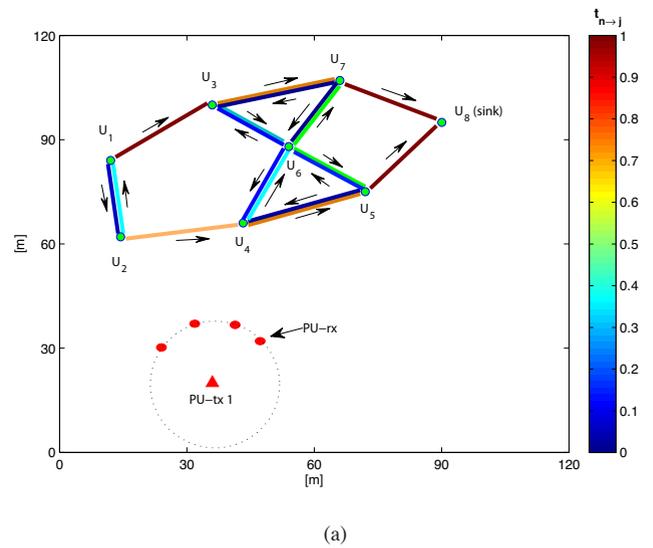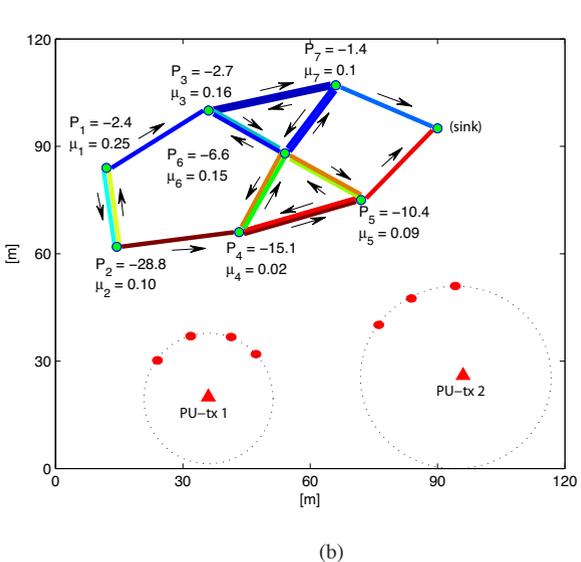
(b)

Fig. 2. Test case 1: routing probabilities $\{t_{n \to i}\}$ (top); and fading-induced outage probabilities (bottom).

(a)

(b)

Fig. 3. Test case 2 with the second PU transmitter inactive: routing probabilities $\{t_{n \to i}\}$ (top); and fading-induced outage probabilities (bottom).

TABLE I
EXOGENOUS TRAFFIC RATES.

	U_1	U_2	U_3	U_4	U_5	U_6	U_7
Test 1	0.052	0.023	0.028	0.01	0.011	0.008	0.044
Test 2	0.057	0.038	0.06	0.05	0.015	0.011	0.051

experiments, are reported in Table I. It can be seen that ρ_2, ρ_4, and ρ_5 increase in this case. This example demonstrates the capability of the proposed routing approach to adapt routes and transmit-powers to locations of active PUs.

Fig. 4 corroborates the convergence of Algorithm 1 for $\beta = 0.1$ and $c \in \{1, 10\}$. Specifically, the depicted evolution of $|t_{n \to j}(k) - t_{n \to j, j}(k)|$ for nodes U_3 and U_6 shows that the local routing probabilities approximately coincide with those of the neighboring nodes after a few iterations. For example, a gap smaller than 1% is obtained after 8 iterations. A similar trend was observed for the transmit-probabilities, which suggests that an online implementation of the algorithm is feasible, and queues will be stable after just a few iterations.

VII. CONCLUSIONS

A novel cross-layer optimization framework was introduced in this paper. Based on channel and interference level statistics, and the situational awareness provided by spectrum sensing schemes, the novel approach yields optimal routes, transmission probabilities, and transmit-powers. The relevant optimization problem turned out to be non-convex and hence difficult to solve even in a centralized setup. Nevertheless, a successive convex approximation was pursued to find a KKT solution. Primal decomposition and AD-MoM were employed to derive a distributed algorithm, suitable for large networks, and amenable to online implementation. As packets are randomly routed through the network, their deliverability in case of time-varying routing strategies and link reliabilities was asserted. Finally, numerical tests verified the ability of the proposed scheme to adapt network operation to the propagation environment.
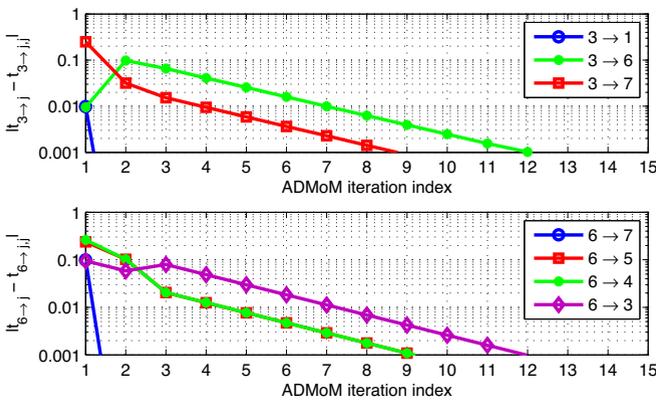
Fig. 4. Convergence of Algorithm 1.

REFERENCES

[1] J. L. Wang and J. A. Silvester, "Maximum number of independent paths and radio connectivity," *IEEE Trans. Wireless Commun.*, vol. 41, no. 10, pp. 1482–1493, Oct. 1993.

[2] M. Haenggi, "On routing in random Rayleigh fading networks," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1553–1562, Jul. 2005.

[3] D. D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *Proc. Int. ACM Conf. Mobile Computing, Networking*, San Diego, CA, Sep. 2003, pp. 134–156.

[4] H. Liu, Z.-L. Zhang, J. Srivastava, and V. Firoiu, "Pwave: A multi-source multi-sink anycast routing framework for wireless sensor networks," in *Proc. Intl. Conf. on Networking*, Atlanta, GA, May 2007.

[5] A. Ribeiro, Z.-Q. Luo, N. Sidiropoulos, and G. B. Giannakis, "Modelling and optimization of stochastic routing for wireless multihop networks," in *Proc. IEEE Int. Conf. on Computer Commun.*, Anchorage, AK, May 2007, pp. 1748–1756.

[6] A. Ephremides, "Energy concerns in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 48–59, Aug. 2002.

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

[8] H. Khalife, S. Ahuja, N. Malouch, and M. M. Krunz, "Probabilistic path selection in opportunistic cognitive radio networks," in *Proc. IEEE Glob. Telecom. Conf.*, New Orleans, LO, Dec. 2010.

[9] C. Xin, L. Ma, and C.-C. Shen, "A path-centric channel assignment framework for cognitive radio wireless networks," *Mobile Net. Appl.*, vol. 13, no. 5, pp. 463–476, Oct. 2008.

[10] I. Pefkianakis, S. Wong, and S. Lu, "SAMER: Spectrum aware mesh routing in cognitive radio networks," in *Proc. IEEE DySPAN*, Chicago, IL, Oct. 2008.

[11] A. Abbagnale and F. Cuomo, "Connectivity-driven routing for cognitive radio ad-hoc networks," in *Proc. IEEE SECON*, Boston, MA, 2010.

[12] K. R. Chowdhury and I. F. Akyildiz, "CRP: A routing protocol for cognitive radio ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 794–802, Apr. 2011.

[13] S.-J. Kim, E. Dall'Anese, and G. B. Giannakis, "Cooperative spectrum sensing for cognitive radios using Kriged Kalman filtering," *IEEE J. Sel. Topics Signal Proc.*, vol. 5, pp. 24–36, Feb. 2011.

[14] E. Dall'Anese, J. A. Bazerque, and G. B. Giannakis, "Group sparse Lasso for cognitive network sensing robust to model uncertainties and outliers," *Elsevier Physical Communication*, Nov. 2011.

[15] C. Lott and D. Teneketzis, "Stochastic routing in ad-hoc networks," *IEEE Trans. Autom Control*, vol. 51, no. 1, pp. 52–70, Jan. 2006.

[16] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Boston, MA: Kluwer Academic Publishers, 2001.

[17] E. Dall'Anese, S.-J. Kim, G. B. Giannakis, and S. Pupolin, "Power control for cognitive radio networks under channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 3541–3551, Dec. 2011.

[18] L. F. Fenton, "The sum of lognormal probability distributions in scatter transmission systems," *IRE Trans. Commun. Syst.*, vol. 8, no. 1, pp. 57–67, Mar. 1960.

[19] S. Shin, S. Choi, H. S. Park, and W. H. Kwon, "Packet error rate analysis of IEEE 802.15.4 under IEEE 802.11b interference," in *Proc. WWIC 2005, LNCS, Springer*, May. 2004, pp. 279–288.

[20] A. A. Abu-Dayya and N. C. Beaulieu, "Comparison of methods of computing correlated lognormal sum distributions and outages for digital wireless applications," in *Proc. IEEE Veh. Tech. Conf.*, May 1994, pp. 175–179.

[21] R. Rao and A. Ephremides, "On the stability of interacting queues in a multi-access system," *IEEE Trans. Inf. Theory*, vol. 34, pp. 918–930, Sep. 1988.

[22] M. J. N. L. Georgiadis and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends in Netw.*, vol. 1, no. 1, pp. 1–144, 2006.

[23] R. Loynes, "The stability of a queue with non-independent interarrival and service times," *Mathematical Proc. the Cambridge Philosophical Society*, vol. 58, pp. 497–520, 1962.

[24] E. Dall'Anese, S.-J. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed Kriging," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1205–1211, Mar. 2011.

[25] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.

[26] A. G. Marques, N. Gatsis, and G. B. Giannakis, "Optimal cross-layer design of wireless fading multi-hop networks," *in Cross Layer Designs in WLAN Systems*, N. Zorba, C. Skianis, and C. Verikoukis, Eds. Leicester, UK: Troubador Pub., 2011.

[27] A. Ribeiro, "Wireless cooperative communications and networking," PhD thesis, University of Minnesota. 2006. [Online]: http://www.seas.upenn.edu/~aribeiro/preprints/.

[28] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 840–845, Jul. 2003.

[29] N. Ermolova and S.-G. Haggman, "Simplified bounds for the complementary error function; application to the performance evaluation of signal processing systems," in *Proc. the 12th European Signal Proces. Conf.*, Vienna, Austria, Sep. 2004.

[30] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, Jul.-Aug. 1978.

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[32] A. Ribeiro, N. Sidiropoulos, and G. B. Giannakis, "Optimal distributed stochastic routing algorithms for wireless multihop networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4261–4272, Nov. 2008.

[33] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 9, pp. 1439–1451, 2006.

[34] A. Nedić and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1757–1780, 2009.

[35] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[36] C. Michelot, "A finite algorithm for finding the projection of a point onto the canonical simplex of $\mathbb{R}^n$," *J. Optim. Theory Appl.*, vol. 50, no. 1, pp. 195–200, Jul. 1986.

[37] M. E. J. Newman, *Networks: An Introduction*. Oxford University Press, 2010.

[38] M. Avriel and A. C. Williams, "An extension of geometric programming with applications in engineering optimization," *J. Engineering Mathematics*, vol. 5, no. 3, pp. 187–194, Jul. 1971.

Emiliano Dall'Anese (S'08, M'11) received the Laurea Triennale (B.Sc degree) and the Laurea Specialistica (M.Sc degree) in Telecommunications Engineering from the University of Padova, Italy, in 2005 and 2007, respectively, and the Ph.D in Information Engineering at the Department of Information Engineering (DEI), University of Padova, Italy, in 2011. From January 2009 to September 2010 he was a visiting scholar at the Department of Electrical and Computer Engineering, University of Minnesota, USA. He is currently a post-doctoral associate at the Department of Electrical and Computer Engineering, University of Minnesota, USA.

His research interests lie in the areas of statistical signal processing, communication theory, and networking. Current research focuses on wireless cognitive radio systems, IP networks, and power distribution networks.

Georgios B. Giannakis (F'97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. Since 1999 he has been a professor with the Univ. of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE Department, and serves as director of the Digital Technology Center.

His general interests span the areas of communications, networking and statistical signal processing - subjects on which he has published more than 300 journal papers, 500 conference papers, 20 book chapters, two edited books and two research monographs. Current research focuses on compressive sensing, cognitive radios, network coding, cross-layer designs, wireless sensors, social and power grid networks. He is the (co-) inventor of twenty patents issued, and the (co-) recipient of eight paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.