Network Integrity in Mobile Robotic Networks

Michael M. Zavlanos, *Member, IEEE*, Alejandro Ribeiro, *Member, IEEE*, and George J. Pappas, *Fellow, IEEE*

***Abstract*—Most coordinated tasks performed by teams of mobile robots require reliable communications between team members. Therefore, task accomplishment requires that robots navigate their environment with their collective movement restricted to formations that guarantee integrity of the communication network. Maintaining this communication capability induces physical constraints on trajectories but also requires determination of communication variables like routes and transmitted powers. This problem is addressed here using a distributed hybrid approach. Continuous distributed motion controllers based on potential fields interact with discrete distributed optimization of communication variables to result in a muti-robot network ensuring communication integrity. The definition of network integrity differs from existing approaches in that it is not based only on the topology of the network but also on metrics that are of interest to the performance of communication between robots and a fixed infrastructure. Specifically, integrity is defined as the ability of the network to support desired communication rates. The ability of the hybrid controller to guarantee communication integrity while robots move to accomplish their task is studied theoretically and numerically.**

***Index Terms*—Convex optimization, distributed algorithms, hybrid systems, network optimization, robotic teams.**

I. Introduction

MOBILE robot networks have recently emerged as an inexpensive and robust way to address a wide variety of tasks ranging from exploration, surveillance and reconnaissance, to cooperative construction and manipulation. Efficient information exchange and coordination between members of the team are critical for successful completion of these tasks. E.g., recent work on distributed consensus and state agreement has strongly depended on multi-hop communication for convergence and performance guarantees [1]–[3].

Multi-hop communication in multi-robot systems has typically relied on constructs from graph theory, with proximity graphs gaining the most popularity. This is consistent with early approaches to wireless networking that used disk models to abstract the physical layer [4]–[9]. In this context, communication becomes equivalent to topological connectivity, defined as the property of a graph to transmit information between all pairs of nodes.

Manuscript received September 01, 2011; revised September 01, 2011; accepted May 09, 2012. The work of M. M. Zavlanos is supported by the NSF CNS Grant #1054604. The work of A. Ribeiro and G. J. Pappas was supported by NSF CPS under Grant 0931239. Date of publication June 06, 2012; date of current version December 17, 2012. Recommended by Associate Editor P. Tabuada.

M. M. Zavlanos is with the Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708 USA (e-mail: michael.zavlanos@duke.edu).

A. Ribeiro and G. J. Pappas are with the Department of Electrical and Systems Engineering, the University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: aribeiro@seas.upenn.edu; pappasg@seas.upenn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2012.2203215

Preservation and control of topological connectivity was first addressed in [10] in the context of connectivity preserving rendezvous. Since then, it has gained increased popularity with approaches that strictly maintain communication links [11]–[14], being followed by least restrictive ones that allow links to be lost [15], [16]. In terms of solution techniques, approaches are both centralized [11], [14] and distributed [12], [13], [15], [16], with the former typically based on semidefinite programming [11], [12], [17] or potential fields [14], and the latter on switched and hybrid systems theory [13], [15], [16].

Although graphs provide a simple abstraction of inter-robot communications, it has long being recognized that since links in a wireless network do not entail tangible connections, associating links with arcs on a graph can be somewhat arbitrary. Indeed, topological definitions of connectivity start by setting target signal strengths to draw the corresponding graph. Even small differences in target strengths might result in dramatic differences in network topology [18]. As a result, graph connectivity is necessary but not nearly sufficient to guarantee communication integrity, interpreted as the ability of a network to support desired communication rates.

This paper employs a simple, yet effective, modification that relies on weighted graph models with weights that capture the packet error probability of each link [19]. When using reliabilities as link metrics it is possible to model routing and scheduling problems as optimization problems that accept link reliabilities as inputs [20], [21]. The key idea proposed in this paper is to define connectivity in terms of communication rates and to use optimization formulations to describe optimal operating points of wireless networks. The use of optimization as a mathematical tool to analyze network protocols dates back to [22] and has been extensively used in wired [23], [24] and wireless networks [25]–[27]. General optimal wireless networking problems are defined to determine end-to-end user rates, routes, link capacities, and transmitted power, as well as frequency and power allocations [28]–[33]. Due to the lack of a duality gap, such problems can be simplified by working in the dual domain [28], [29], [31]–[33].

The main contribution of this work is the use of optimal wireless network design to develop novel alternatives for mobility control. In particular, we decompose control in the communication and physical domains, so that the communication variables are updated in discrete time via a distributed gradient descent algorithm on the dual function, while robot motion is regulated in continuous time by means of appropriate distributed barrier potentials that maintain desired communication rates. Composition of these techniques results in a distributed multi-robot hybrid system for which we show that desired communication rates are always guaranteed. The challenge we need to address is that working in the dual domain ensures feasibility of the primal variables only asymptotically. In fact, as the robots move, the optimal solution in the communications space drifts, which introduces an

infeasibility gap in the primal variables. This precludes verbatim use of those variables in barrier potentials in the physical domain. We provide bounds for this gap in terms of problem specific parameters and propose mobility controllers that ensure stability of the integrated system, in that desired bounds on the infeasibility gap are always met and a local optimum in the joint communication and mobility space is reached. A related problem that considers optimal communications based on T-slot time averages of the primal variables for general mobility schemes was recently addressed in [34]. The difference with our work, is that [34] does not explicitly control robot motion.

The rest of this paper is organized as follows. Section II defines network integrity in terms of communication rates and develops a dual decomposition for distributed optimal communication. Section III integrates distributed motion control with optimization of the communication variables in a distributed hybrid multi-robot system. Section IV discusses performance of the proposed algorithm. Section V presents simulations to illustrate the performance of the proposed approach.

II. Optimal Wireless Networking

Consider a mobile network composed of J robots and a fixed infrastructure with K access points (APs). The robots move throughout an area of interest to accomplish an assigned task for which it is necessary to maintain reliable communications with the infrastructure. Due to, e.g., power constraints or an adverse propagation environment, robots collaborate to maintain a multihop network with the APs.

Let $\mathbf{x}_i$ for $i = 1, \ldots, J$ denote the positions of the robots and $\mathbf{x}_i$ for $i = J+1, \ldots, J+K$ the positions of the APs. The set of all positions $\mathbf{x} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_{J+K}^T]^T$ is referred to as the spatial configuration of the network. We model communication by a link reliability metric $R(\mathbf{y}_1, \mathbf{y}_2)$ denoting the probability that a packet transmitted from a terminal located at position $\mathbf{y}_1$ is correctly decoded by a terminal at position $\mathbf{y}_2$. This function determines the probability $R_{ij} \triangleq R(\mathbf{x}_i, \mathbf{x}_j)$ with which a packet transmitted by node i is correctly decoded by node j. Node j is a robot if $j \leq J$ or an AP otherwise. Using R_0 to denote the transmission rate of the terminals' radios, the effective transmission rate from i to j is the rate $R_0 R(\mathbf{x}_i, \mathbf{x}_j)$ at which information is successfully conveyed in this link. To simplify notation we work with normalized rates by making $R_0 = 1$. This means that rates are measured as (dimensionless) fractions of the transmission rate R_0. To recover rates measured in information units per second it suffices to multiply by R_0. Here, we assume all robots use the same transmission rate R_0. This assumption is easy to lift.

We further denote as $r_i \in [0, 1]$ the normalized average rate at which robot i delivers information to the APs. If robot i can reach some of the APs, which is possible if the probability $R(\mathbf{x}_i, \mathbf{x}_j)$ is reasonably large for some $j \in \{J+1, \ldots, J+K\}$, packets are directly conveyed to the corresponding AP. Otherwise, packets are routed to another robot for subsequent transmission. We model this process through the introduction of routing probabilities T_{ij} denoting the probability with which robot i selects node j, either a robot or an AP, as a destination of its transmitted packets; see Fig. 1.

Between their generation or arrival from another robot and their transmission, packets are stored in a queue; see Fig. 2.

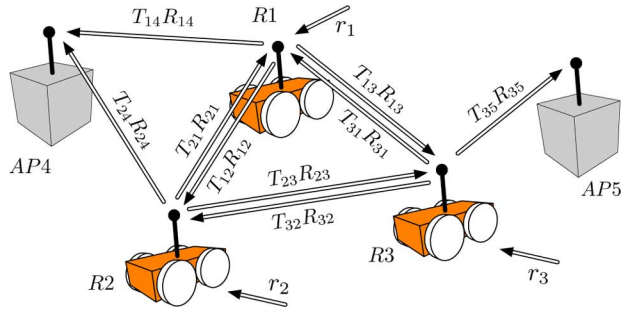


Fig. 1. Robot network consisting of two access points (APs) and three robots (Rs). Robots generate data packets related to their assigned task at a rate $0 \leq r_i \leq 1$. They collaborate to convey this information to the APs by relaying data for each other. The position function $R(\mathbf{x}_i, \mathbf{x}_j)$ is the reliability of the channel between robots i and j, whereas T_{ij} denotes the probability that robot i routes packets to robot j. The product $T_{ij}R(\mathbf{x}_i, \mathbf{x}_j)$ is the average rate at which packets are successfully conveyed from robot i to terminal j.

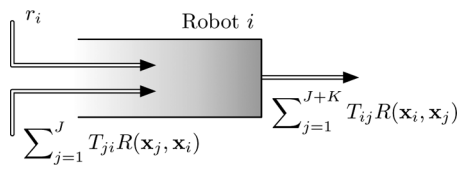


Fig. 2. Queue at robot i.

Packets leave the queue at robot i upon transmittion to any other node j and successful decoding by this intended next-hop. Since these two events are independent, the normalized rate at which packets are sent from i to j is $T_{ij}R(\mathbf{x}_i, \mathbf{x}_j)$. The aggregate rate at which packets leave the i-th queue is then

$$r_i^{\text{out}} = \sum_{j=1}^{J+K} T_{ij}R(\mathbf{x}_i, \mathbf{x}_j). \tag{1}$$

Similarly, a packet arrives at robot i coming from robot j, when robot j selects i as the next hop and i correctly decodes the packet. This happens with probability $T_{ji}R(\mathbf{x}_j, \mathbf{x}_i)$. Considering that packets are also locally generated at a rate r_i, the rate at which packets arrive at the i-th queue is

$$r_i^{\text{in}} = r_i + \sum_{j=1}^{J} T_{ji}R(\mathbf{x}_j, \mathbf{x}_i). \tag{2}$$

Note that the sum in (1) is up to $J+K$ because packets can be sent to another robot or an AP, whereas the sum in (2) is up to J because packets are received from peer robots only.

If the average rate at which packets arrive at the i-th queue is smaller than the average rate at which packets leave this queue, i.e., if $r_i^{\text{in}} \leq r_i^{\text{out}}$, then the queue empties infinitely often with probability one. This provides an almost sure guarantee that packets are eventually delivered to the AP as long as $r_i^{\text{in}} \leq r_i^{\text{out}}$. Thus, our interest is to determine routing probabilities T_{ij} and rates r_i that satisfy the flow inequality

$$r_i \leq \sum_{j=1}^{J+K} T_{ij}R(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^{J} T_{ji}R(\mathbf{x}_j, \mathbf{x}_i). \tag{3}$$

Any set of variables $\{r_i\}_{\forall\, i}$ and $\{T_{ij}\}_{\forall\, i,j}$ that satisfy (3) ensures information delivery. A basic requirement is that all robots communicate with the infrastructure APs at least at a basal rate of r_{i0} packets per time unit. When this happens we say that we have network integrity as we formally define next.

Definition 1 (Network Integrity): Given positions $\mathbf{x}_i$ and basal rates r_{i0}, network integrity implies the availability of rates $\{r_i\}_{\forall\, i}$ and routing variables $\{T_{ij}\}_{\forall\, i,j}$ for which the inequalities in (3) hold and $r_i \geq r_{i0}$, for all i.

We remark that network integrity refers not only to existence but availability of variables satisfying the stated constraints. For a given spatial configuration $\mathbf{x}$ there might be various sets of variables that ensure network integrity. To select an element of this set we introduce strictly concave optimality criteria $U_i(r_i)$ and $V_{ij}(T_{ij})$ measuring the value associated with variables r_i and T_{ij} respectively. The operating point is then selected as the solution of the optimization problem

$$\begin{aligned} P_{\mathbf{x}} = \max_{r_i,\, T_{ij}} \quad & \sum_{i=1}^{J} U_i(r_i) + \sum_{i=1}^{J}\sum_{j=1}^{J+K} V_{ij}(T_{ij}) \\ \text{s.t} \quad & r_i + \sum_{j=1}^{J} T_{ji} R(\mathbf{x}_j, \mathbf{x}_i) \leq \sum_{j=1}^{J+K} T_{ij} R(\mathbf{x}_i, \mathbf{x}_j), \\ & r_i \geq r_{i0}, \quad \sum_{j=1}^{J+K} T_{ij} \leq 1, \end{aligned} \tag{4}$$

where the flow inequality constraint [cf. (3)], the basal rate constraint $r_i \geq r_{i0}$, and the probability constraint $\sum_{j=1}^{J+K} T_{ij} \leq 1$ are required for all $i \in \{1, \ldots, J\}$. To ensure network integrity for given robot positions $\mathbf{x}$, we need to find optimal routing probabilities $T^*_{\mathbf{x},ij}$ and rates $r^*_{\mathbf{x},i}$ that solve the optimization problem in (4). This implies rates $r^*_{\mathbf{x},i}$ that exceed desired basal rates r_{i0} for all terminals, while assigning the remaining resources in a manner that is optimal in terms of the utilities $U_i(r_i)$ and $V_{ij}(T_{ij})$. Observe that the optimal utility yield $P_{\mathbf{x}}$ depends on the spatial configuration $\mathbf{x}$, as so do the sets of optimal routing and rate variables $T^*_{\mathbf{x},ij}$ and $r^*_{\mathbf{x},i}$. For fixed spatial configuration $\mathbf{x}$, the reliabilities $R(\mathbf{x}_i, \mathbf{x}_j)$ are fixed and the problem in (4) attains a simple convex form.

Remark 1 (Utility Functions): Utilities $U_i(r_i)$ in (4) are metrics that compare different network operating points. Thinking of nodes as economic agents and of utilities $U_i(r_i)$ as the value of rate r_i, the network's objective is to maximize the social value $\sum_i U_i(r_i)$. Example utilities are linear $U_i(r_i) = w_i r_i$ and logarithmic $U_i(r_i) = w_i \log(r_i)$. Linear utilities yield larger rates, while logarithmic utilities yield fairer operating points because they penalize small rates r_i. In most network optimization problems utilities $V_{ij}(T_{ij})$ associated with routing variables T_{ij} are not used, i.e., they are set to $V_{ij}(T_{ij}) = 0$. In the problems considered here $V_{ij}(T_{ij})$ may be useful. E.g., making $V_{ij}(T_{ij}) = -w_{ij} T_{ij}^2$ encourages packet splitting through different paths. Packet splitting improves robustness to link and robot failures and increases the ability of the algorithm to discover neighboring robots. The required strict concavity of $V_{ij}(T_{ij})$ precludes the use of $V_{ij}(T_{ij}) = 0$ in (4).

A. Distributed Optimal Communication

Solving (4) necessitates discovery of the network's topology by a designated terminal, followed by dissemination of the optimal operating variables. This process not only entails a large communication cost but it also incurs significant delays and is vulnerable to failures. Alternatively, it is possible to devise a distributed solution by exploiting the separability of the Lagrangian dual of (4). To do so, consider the optimal communication problem in (4), associate Lagrange multipliers λ_i with each of the routing constraints in (3), and define the Lagrangian as

$$\begin{aligned} \mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda}, \mathbf{T}, \mathbf{r}) &= \sum_{i=1}^{J} U_i(r_i) + \sum_{i=1}^{J}\sum_{j=1}^{J+K} V_{ij}(T_{ij}) \\ &+ \sum_{i=1}^{J} \lambda_i \Big[\sum_{j=1}^{J+K} T_{ij} R(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^{J} T_{ji} R(\mathbf{x}_j, \mathbf{x}_i) - r_i \Big], \end{aligned} \tag{5}$$

where $\boldsymbol{\lambda}$, $\mathbf{r} \in \mathbb{R}^J$ are column vectors of Lagrange multipliers λ_i and rates r_i, respectively, and $\mathbf{T} \in \mathbb{R}^{J \times J+K}$ is a matrix of routing probabilities T_{ij}. The dual function is then defined as the maximum of the Lagrangian with respect to primal variables that satisfy the constraints that were not included in the Lagrangian, i.e.,

$$g_{\mathbf{x}}(\boldsymbol{\lambda}) = \max_{r_i \geq r_{i0},\, \sum_{j=1}^{J+K} T_{ij} \leq 1} \mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda}, \mathbf{T}, \mathbf{r}), \tag{6}$$

where the constraints $r_i \geq r_{i0}$ and $\sum_{j=1}^{J+K} T_{ij} \leq 1$ are required for all i. The dual problem is finally defined as the minimization of the dual function, $D_{\mathbf{x}} = \min_{\boldsymbol{\lambda} \geq \mathbf{0}} g_{\mathbf{x}}(\boldsymbol{\lambda})$, with the corresponding set of optimal dual variables denoted as $\boldsymbol{\lambda}^*_{\mathbf{x}}$. Notice that the Lagrangian $\mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda}, \mathbf{T}, \mathbf{r})$, the dual function $g_{\mathbf{x}}(\boldsymbol{\lambda})$, and the optimal dual value $D_{\mathbf{x}}$ are all defined parametrically with respect to spatial configuration $\mathbf{x}$. Since for given positions $\mathbf{x}$, the problem in (4) is convex it holds that $D_{\mathbf{x}} = P_{\mathbf{x}}$ implying that we can work with the dual problem in lieu of the primal problem in (4). In particular, a distributed algorithm can be obtained by implementing gradient descent in the dual domain.

To implement gradient descent we need to compute a gradient of the dual function, which can be obtained from the primal Lagrangian maximizers, see e.g., [35]. For given $\boldsymbol{\lambda}$ and spatial configuration $\mathbf{x}$ define the primal Lagrangian maximizers as

$$\{r_{\mathbf{x},i}(\boldsymbol{\lambda})\}_{\forall\, i}, \{T_{\mathbf{x},ij}(\boldsymbol{\lambda})\}_{\forall\, i,j} \triangleq \operatorname*{argmax}_{\substack{r_i \geq r_{i0}, \\ \sum_{j=1}^{J+K} T_{ij} \leq 1}} \mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda}, \mathbf{T}, \mathbf{r}). \tag{7}$$

The components of the dual function's gradient are then given by the constraint slack associated with $\{r_{\mathbf{x},i}(\boldsymbol{\lambda})\}_{\forall\, i}$ and $\{T_{\mathbf{x},ij}(\boldsymbol{\lambda})\}_{\forall\, i,j}$, i.e.,

$$\begin{aligned} [\nabla g_{\mathbf{x}}(\boldsymbol{\lambda})]_i = & \sum_{j=1}^{J+K} T_{\mathbf{x},ij}(\boldsymbol{\lambda}) R(\mathbf{x}_i, \mathbf{x}_j) \\ & - \sum_{j=1}^{J} T_{\mathbf{x},ji}(\boldsymbol{\lambda}) R(\mathbf{x}_j, \mathbf{x}_i) - r_{\mathbf{x},i}(\boldsymbol{\lambda}). \end{aligned} \tag{8}$$

A key observation here is that the Lagrangian in (5) can be written as a sum of local Lagrangians that depend only on variables r_i and $\{T_{ij}\}_{\forall\, i}$. Indeed, it suffices to reorder terms in (5) to realize that upon defining local Lagrangians

$$\begin{aligned} \mathcal{L}_{\mathbf{x},i}(\boldsymbol{\lambda}, \mathbf{T}, \mathbf{r}) &= U_i(r_i) - \lambda_i r_i \\ &+ \sum_{j=1}^{J} \Big[V_{ij}(T_{ij}) + T_{ij} R(\mathbf{x}_i, \mathbf{x}_j)(\lambda_i - \lambda_j) \Big] \\ &+ \sum_{j=J+1}^{J+K} \Big[V_{ij}(T_{ij}) + \lambda_i T_{ij} R(\mathbf{x}_i, \mathbf{x}_j) \Big] \end{aligned} \tag{9}$$

it is possible to write

$$\mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r})=\sum_{i=1}^{J}\mathcal{L}_{\mathbf{x},i}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r}). \tag{10}$$

The local Lagrangian $\mathcal{L}_{\mathbf{x},i}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r})$ is defined so that all summands of the global Lagrangian $\mathcal{L}_{\mathbf{x}}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r})$ that involve primal variables r_i and $\{T_{ij}\}_{j=1}^{J+K}$ for given i appear in, and only in, $\mathcal{L}_{\mathbf{x},i}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r})$ [cf. (5) and (9)]. Therefore, to find the variables $r_{\mathbf{x},i}(\boldsymbol{\lambda})$ and $\{T_{\mathbf{x},ij}(\boldsymbol{\lambda})\}_{j=1}^{J+K}$ that maximize the global Lagrangian as per (7) it suffices to find the arguments that maximize the local Lagrangians in (9),

$$r_{\mathbf{x},i}(\boldsymbol{\lambda}),\{T_{\mathbf{x},ij}(\boldsymbol{\lambda})\}_{j=1}^{J+K}=\underset{\substack{r_i\geq r_{i0},\\ \sum_{j=1}^{J+K}T_{ij}\leq 1}}{\operatorname{argmax}}\ \mathcal{L}_{\mathbf{x},i}(\boldsymbol{\lambda},\mathbf{T},\mathbf{r}). \tag{11}$$

where the constraints $r_i\geq r_{i0}$ and $\sum_{j=1}^{J+K}T_{ij}\leq 1$ in (11) are for the node i under consideration. Contrast (7) and (11) to observe that in (7) we maximize the global Lagrangian subject to global constraints, while in (11) we maximize local Lagrangians with respect to local constraints.

Introduce now an index n and consider time instants $\{t_n\}_{n=0}^{\infty}$ at which variables are updated. We can use the observation in (11) to write the following distributed gradient descent algorithm for the dual function:

1) Primal Iteration: For given Lagrange multipliers $\boldsymbol{\lambda}(t_n)$ and spatial configuration $\mathbf{x}(t_n)$ compute Lagrangian maximizers $r_i(t_n)=r_{\mathbf{x}(t_n),i}[\boldsymbol{\lambda}(t_n)]$ and $T_{ij}(t_n)=T_{\mathbf{x}(t_n),ij}[\boldsymbol{\lambda}(t_n)]$, where $r_{\mathbf{x}(t_n),i}[\boldsymbol{\lambda}(t_n)]$ and $T_{\mathbf{x}(t_n),ij}[\boldsymbol{\lambda}(t_n)]$ are as defined in (7) for $\mathbf{x}=\mathbf{x}(t_n)$ and $\boldsymbol{\lambda}=\boldsymbol{\lambda}(t_n)$. According to (11), $r_i(t_n)$ and $\{T_{ij}(t_n)\}_{j=1}^{J+K}$ can be computed as

$$r_i(t_n),\{T_{ij}(t_n)\}_{j=1}^{J+K}=\underset{\substack{r_i\geq r_{i0},\\ \sum_{j=1}^{J+K}T_{ij}\leq 1}}{\operatorname{argmax}}\ \mathcal{L}_{\mathbf{x}(t_n),i}(\boldsymbol{\lambda}(t_n),\mathbf{T},\mathbf{r}). \tag{12}$$

2) Dual Iteration: Use the primal variables $r_i(t_n)$ and $T_{ij}(t_n)$ in (12) to update the dual variables as

$$\lambda_i(t_{n+1})=\mathbb{P}\Big[\lambda_i(t_n)-\epsilon\Big(\sum_{j=1}^{J+K}T_{ij}(t_n)R(\mathbf{x}_i(t_n),\mathbf{x}_j(t_n)) -\sum_{j=1}^{J}T_{ji}(t_n)R(\mathbf{x}_j(t_n),\mathbf{x}_i(t_n))-r_i(t_n)\Big)\Big] \tag{13}$$

where $\mathbb{P}[\cdot]$ denotes the projection on the nonnegative orthant.

Letting variables $r_i(t_n)$, $\{T_{ij}(t_n)\}_{j=1}^{J+K}$, and $\lambda_i(t_n)$ be associated with terminal i, the algorithm in (12)–(13) can be implemented in a distributed manner. The maximization in (12) requires access to local multipliers $\lambda_i(t_n)$ and multipliers $\lambda_j(t_n)$ from neighboring terminals for which $R(\mathbf{x}_j(t_n),\mathbf{x}_i(t_n))\neq 0$. Since these terminals can communicate with i, these multipliers can be conveyed to let i compute its primal variables of interest. Likewise, the dual update in (13) requires access to local primal variables $r_i(t_n)$ and $\{T_{ij}(t_n)\}_{j=1}^{J+K}$ as well as neighboring primal variables $\{T_{ji}(t_n)\}_{j=1}^{J}$ from terminals that can communicate directly with i.

In (13), the Lagrange multiplier $\lambda_i(t_n)$ is updated as determined by the i-th component $\big[\nabla g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n))\big]_i$ of the gradient of the dual function corresponding to the spatial configuration $\mathbf{x}(t_n)$ at the time of the update [cf. (8) and (13)]. On aggregate, the updates in (13) define a projected—recall that $\lambda_i\geq 0$—gradient descent step on the dual function $g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n))$. If spatial configurations are fixed, i.e., if $\mathbf{x}(t_n)=\mathbf{x}$ for all n, we expect that as time progresses multiplier iterates $\boldsymbol{\lambda}(t_n)$ approach the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}}$. Provided some technical conditions that we discuss in Section IV, convergence of the primal variables $T_{ij}(t_n)$ and $r_i(t_n)$ to the optimal network operating point $T^*_{\mathbf{x},ij}, r^*_{\mathbf{x},i}$ follows as a consequence. The effect of time varying spatial configurations on the performance of (12)–(13) is analyzed in Section IV.

III. Distributed Motion & Communication Control

In Section II, we showed that for fixed robot positions $\mathbf{x}$, the reliabilities $R(\mathbf{x}_i,\mathbf{x}_j)$ are fixed and the problem in (4) attains a simple convex form. However, this is not the case for mobile robots that are required to move to accomplish their assigned task. To address this challenge, we decompose control in the motion and communication domains so that the communication variables are updated at discrete time instances $\{t_n\}_{n=0}^{\infty}$ via the distributed primal-dual iteration (12)–(13) while robot positions $\mathbf{x}_i(t)$ for $i=1,\dots,J$ are updated in continuous time. Between times t_n and t_{n+1} positions $\mathbf{x}_i(t)$ are updated according to the first order differential equations

$$\dot{\mathbf{x}}_i(t)=\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n)),\quad \forall\, t\in[t_n,t_{n+1}), \tag{14}$$

where the function $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ denotes the control signal of robot i and $\sigma_i(t_n)$ denotes a switching signal defined as

$$\sigma_i(t_n)\triangleq\{T_{ij}(t_n)\}_{j=1}^{J+K}\cup\{T_{ji}(t_n)\}_{j=1}^{J}. \tag{15}$$

The switching signal $\sigma_i(t_n)$ contains all the routing variables $\{T_{ij}(t_n)\}_{j=1}^{J+K}$ for traffic out of i and all the variables $\{T_{ji}(t_n)\}_{j=1}^{J}$ for traffic incoming to i. Routing variables $\{T_{ij}(t_n)\}_{j=1}^{J+K}$ are locally available at robot i whereas routing variables $\{T_{ji}(t_n)\}_{j=1}^{J}$ are received through communication with neighboring terminals, i.e., terminals for which $R(\mathbf{x}_j(t_n),\mathbf{x}_i(t_n))\neq 0$. The collection of all hybrid robots, gives rise to a distributed multi-robot hybrid system, where every robot makes local mobility decisions based on information received from neighboring peers.

The definition of $\sigma_i(t_n)$ is such that it contains all the information that allows us to compute the offered rate at time t which we define as

$$\beta_i(\mathbf{x}(t);\sigma_i(t_n))\triangleq\sum_{j=1}^{J+K}T_{ij}(t_n)R(\mathbf{x}_i(t),\mathbf{x}_j(t)) -\sum_{j=1}^{J}T_{ji}(t_n)R(\mathbf{x}_j(t),\mathbf{x}_i(t)). \tag{16}$$

The offered rate $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ is the maximum rate that ensures information delivery for robot i when the spatial configuration is $\mathbf{x}(t)$ and robots route their packets according to $\{T_{ij}(t_n)\}_{j=1}^{J+K}$.

Observe the distinction between basal rates r_{i0}, rates $r_i(t_n)$, and offered rates $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$. Basal rate r_{i0} is a minimum system requirement. Rates $r_i(t_n)$ denote the solution to (11) which is chosen as a function of dual variables and exceeds r_{i0} for all t_n. Offered rate $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$

is an upper limit on achievable rates at time t using available routing variables $T_{ij}(t_n)$. In a static setting we are guaranteed that for large enough time index t_n, it holds $\beta_i(\mathbf{x}(t_n);\sigma_i(t_n)) \geq r_i(t_n) \geq r_{i0}$. In a dynamic setting this is not necessarily true because routing variables $T_{ij}(t_n)$ are computed for configuration $\mathbf{x}(t_n)$, not $\mathbf{x}(t)$. In this case we want to monitor the offered rates $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ to ensure that they do exceed the basal requirements r_{i0}. In the language of Definition 1 we want to satisfy network integrity at time t using routing variables $T_{ij}(t_n)$ and rates $r_i = \beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0}$. The rate $r_i(t_n)$ is just an auxiliary variable. It plays no role on how packets are moved through the network except for its effect on the multiplier $\lambda_i(t_{n+1})$ according to the dual update in (12). Offering the network integrity guarantee $\beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0}$ is the design specification for the controller $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ as we summarize in the following problem statement.

Problem 1: Given a sequence of switching signals $\{\sigma_i(t_n)\}_{n=0}^{\infty}$ as defined in (15) with routing variables $\{T_{ij}(t_n)\}_{n=0}^{\infty}$ determined by the primal-dual iteration (12)–(13), determine a set of distributed motion controllers $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ so that for all times t and robots i

$$\beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0}, \tag{17}$$

with $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ as defined in (16). The condition in (17) guarantees network integrity in that offered communication rates $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ exceed basal rate requirements r_{i0} along the trajectories of the closed loop system (14) at all times.

Problem 1 requires determination of routing variables $T_{ij}(t_n)$ and robot trajectories $\mathbf{x}_i(t)$ satisfying network integrity constraints $\beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0}$ for all times $t \geq t_0$. However, the distributed primal-dual iteration (12)–(13) only ensures feasibility of primal variables in the limit of a static system. Assuming fixed positions and $t_n \to \infty$ iterative application of (12)–(13) ensures feasibility of primal iterates $r_i(t_n)$ and $T_{ij}(t_n)$, which in turns guarantees network integrity at time $t = t_n$; i.e., $\beta_i(\mathbf{x}(t_n);\sigma_i(t_n)) \geq r_i(t_n) \geq r_{i0}$. However, given time t_n, primal iterates $r_i(t_n)$ and $T_{ij}(t_n)$ are close to feasible but not necessarily so; i.e., there exists a small error term e for which $\beta_i(\mathbf{x}(t_n);\sigma_i(t_n)) \geq r_i(t_n) - e \geq r_{i0} - e$. As the system moves from $\mathbf{x}(t_n)$ to $\mathbf{x}(t)$ this error may get larger because routes $T_{ij}(t_n)$ were computed for configuration $\mathbf{x}(t_n)$, not for the current positions $\mathbf{x}(t)$. This observation motivates a reformulation of Problem 1, where network integrity constraints $\beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0}$ are uniformly satisfied with some bounded error $e > 0$ as stated next.

Problem 2: Given a sequence of switching signals $\{\sigma_i(t_n)\}_{n=0}^{\infty}$ as defined in (15) with routing variables $\{T_{ij}(t_n)\}_{n=0}^{\infty}$ determined by the primal-dual iteration (12)–(13), determine a set of distributed motion controllers $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ so that for all times t and robots i

$$\beta_i(\mathbf{x}(t);\sigma_i(t_n)) \geq r_{i0} - e, \tag{18}$$

with $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ as defined in (16) and e a prescribed tolerance. The condition in (18) guarantees approximate network integrity. Offered communication rates $\beta_i(\mathbf{x}(t);\sigma_i(t_n))$ achieve or exceed basal rates r_{i0} discounted by tolerance e along trajectories of the closed loop system (14) at all times.

In Section IV we characterize the uniform error $e > 0$ in terms of system-specific parameters and show that it can be made arbitrarily small by reducing the elapsed time $t_{n+1} - t_n$ between communication variables updates or the robots speed; see Theorem 2.

Since $e > 0$ is a uniform error in the network integrity constraints, it is sufficient to design distributed controllers $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ for Problem 2 that ensure satisfaction of the constraints (18) for all time $t \in [t_n, t_{n+1})$. For this, we rely on artificial potential functions that treat violation of the approximate network integrity constraints in (18) as an obstacle in the free configuration space. These are selected as barrier potentials that grow unbounded when the constraints tend to become violated. Then, defining the controllers $\mathbf{u}_i(\mathbf{x}(t);\sigma_i(t_n))$ to be the negative gradient of these potentials ensures that (18) are satisfied for all times $t \in [t_n, t_{n+1})$.

To do so, introduce the local potential function $\gamma_i : \mathbb{R}^2 \to \mathbb{R}_+$ that robot i strives to minimize; see Section V for specific examples. If a robot has no specific task and its sole purpose is to provide communication support we say that it is a relay robot and adopt the convention $\gamma_i(\mathbf{x}_i) = 1$ for all such $\mathbf{x}_i$. Otherwise, we say that the robot is a leader. Further introduce the potential $\tilde{\beta}_i : \mathbb{R}^{Jd} \to \mathbb{R}_+$ defined as

$$\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t_n)) \triangleq \beta_i(\mathbf{x}(t);\sigma_i(t_n)) - r_{i0} + e. \tag{19}$$

The value of $\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t_n))$ measures satisfaction of the approximate network integrity constraint in (18) for the communication rate offered to robot i. The potential $\tilde{\beta}$ can be computed locally because it involves variables σ_i that are either local to robot i or are communicated from neighboring agents defined as those robots for which $R(\mathbf{x}_j(t_n), \mathbf{x}_i(t_n)) \neq 0$.

The task potential γ_i and the network integrity potential $\tilde{\beta}_i$ may conflict. We balance this differing purposes with an artificial potential function $\hat{\phi}_i : \mathbb{R}^{Jd} \to \mathbb{R}_+$ that we define as

$$\hat{\phi}_i(\mathbf{x}(t);\sigma_i(t_n)) \triangleq \frac{\gamma_i^k(\mathbf{x}_i(t))}{\tilde{\beta}_i^2(\mathbf{x}(t);\sigma_i(t_n))},$$

where $k > 0$ is a positive constant. In the proposed construction, γ_i serves as a goal potential to generate a force attracting robots to their tasks and $\tilde{\beta}_i$ serves as an obstacle barrier potential to generate a force to repel the robots from the space of obstacles, defined as the set of positions $\mathbf{x}(t)$ that violate the approximate network integrity constraints (18). In particular, we can define the obstacle-free space for the team of robots as the set

$$\mathcal{F}(\sigma(t_n)) = \{\mathbf{x} \mid \tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t_n)) > 0,\ \forall\, i\}, \tag{20}$$

where $\sigma(t) = \{\sigma_1(t), \ldots, \sigma_J(t)\}$ denotes the collection of switching signals for all robots.

Since $\hat{\phi}_i$ can grow unbounded as the approximate network integrity constraints tend to become violated, i.e., as $\tilde{\beta}_i \to 0$, resulting in unbounded robot speeds, we further introduce a diffeomorphism $\psi : [0, \infty] \to [0, 1]$ with $\psi(y) = y/(1 + y)$ that squashes the image of $\hat{\phi}_i$ from $[0, \infty]$ to $[0, 1]$. Moreover, define the function $\chi(y) = y^{1/k}$ to restrict the effect of the obstacles close to the boundary of the free space

$$\partial\mathcal{F}(\sigma(t_n)) = \{\mathbf{x} \mid \exists\, i, \text{ s.t. } \tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t_n)) = 0\}. \tag{21}$$

Composition of χ, ψ and $\hat{\phi}_i$ results in the artificial potential $\phi_i : \mathbb{R}^{Jd} \to [0,1]$ for every robot i with

$$\phi_i = \chi \circ \psi \circ \hat{\phi}_i = \frac{\gamma_i}{(\gamma_i^k + \tilde{\beta}_i^2)^{1/k}}. \tag{22}$$

The construction of the potential (22) is inspired by the concept of a navigation function [36], [37]. It is shown in [36], [37] that since χ and ψ are diffeomorphisms, ϕ_i and $\hat{\phi}_i$ have the same critical points. It is also shown in [36], [37] that for sufficiently large

Algorithm 1 Integrated control loop at robot i

Require: Initial positions $\mathbf{x}(t_0)$. Initial multipliers $\boldsymbol{\lambda}(t_0)$;

1: for $n = 0$ to ∞ do {repeat for the life of the system}

2: Determine reliabilities $R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))$ and $R(\mathbf{x}_j(t_n), \mathbf{x}_i(t_n))$;

3: Update transmission rates and routing variables as

$$r_i(t_n), \{T_{ij}(t_n)\}_{j=1}^{J+K} = \underset{\substack{r_i \geq r_{i0}, \\ \sum_{j=1}^{J+K} T_{ij} \leq 1}}{\operatorname{argmax}} \mathcal{L}_{\mathbf{x}(t_n),i}(\boldsymbol{\lambda}(t_n), \mathbf{T}, \mathbf{r}).$$

and transmit $\{T_{ij}(t_n)\}_{j=1}^{J+K}$ to neighbors;

4: Compute the local gradient of dual function by

$$\big[\nabla g(\boldsymbol{\lambda}(t_n))\big]_i = \sum_{j=1}^{J+K} T_{ij}(t_n) R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n)) - \sum_{j=1}^{J} T_{ji}(t_n) R(\mathbf{x}_j(t_n), \mathbf{x}_i(t_n)) - r_i(t_n)$$

5: Update Lagrange multiplier by

$$\lambda_i(t_{n+1}) = \mathbb{P}\left[\lambda_i(t_n) - \epsilon \big[\nabla g(\boldsymbol{\lambda}(t_n))\big]_i\right]$$

and transmit $\lambda_i(t_{n+1})$ to neighbors;

6: Continuously update robot position $\mathbf{x}_i(t)$ until time t_{n+1} by

$$\mathbf{x}_i(t) = \mathbf{x}_i(t_n) - \alpha \int_{t_n}^{t} \nabla_{\mathbf{x}_i} \phi_i(\mathbf{x}(\tau); \sigma_i(t_n)) d\tau$$

and transmit $\mathbf{x}_i(t_{n+1})$ to neighbors;

7: **end for**

constant $k > 0$, the effect of the obstacles becomes localized to the boundary of the free space defined in (21). This facilitates robots to achieve their assigned tasks, assuming these tasks are strictly contained in the free space defined in (20).

The control law for every robot i, whether a leader or a relay, can be defined by the negative gradient of the potential ϕ_i in (22), i.e.,

$$\mathbf{u}_i(\mathbf{x}(t); \sigma_i(t_n)) = -\alpha \nabla_{\mathbf{x}_i} \phi_i(\mathbf{x}(t); \sigma_i(t_n)), \tag{23}$$

for all $t \in [t_n, t_{n+1})$, where $\alpha > 0$ is a positive gain to regulate the robot speed. From (14) and (23) we obtain the distributed closed loop system

$$\dot{\mathbf{x}}_i(t) = -\alpha \nabla_{\mathbf{x}_i} \phi_i(\mathbf{x}(t); \sigma_i(t_n)), \tag{24}$$

for all robots $i = 1, \ldots, J$ and all $t \in [t_n, t_{n+1})$.

Integration of the distributed closed loop system (24) for motion control with the distributed primal-dual iterations (12)–(13) for optimal communication gives rise to the distributed motion and communication control algorithm shown in Algorithm 1. In Algorithm 1, lines 3 to 5 correspond to the primal-dual iteration (12)–(13) that updates the communication variables, and line 6 corresponds to an update in the robot positions via the closed loop system (24). At the time instances $\{t_n\}_{n=0}^{\infty}$ communication between neighbors also takes place to provide the robots with those variables that are necessary in their updates. Correctness of Algorithm 1 relies on showing that a bound on the error $e > 0$ exists, as claimed earlier in this section, and that with this bound, the closed loop system (24) ensures satisfaction of (18) for all time. This analysis is the goal of Section IV.

Remark 2 (Initialization): Our controllers implicitly assume that at time t_0 variables $\mathbf{r}(t_0)$ and $\mathbf{T}(t_0)$ are almost primal feasible in the sense that they satisfy (18). This can be achieved by, e.g., starting with an initialization phase in which the primal-dual iterations in (12)–(13) are run until the flow constraint is satisfied within uniform e tolerance. Alternatively, it is possible to initialize the team in a simple spatial configuration with all robots close to each other so that at time t_0 optimal communication variables are easy to find and we can start with $\boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}^*$, $\mathbf{r}(t_0) = \mathbf{r}^*_{\mathbf{x}(t_0)}$, and $\mathbf{T}(t_0) = \mathbf{T}^*_{\mathbf{x}(t_0)}$.

Remark 3 (Task Completion): In the proposed approach, task completion is a secondary objective, subsidiary to communication integrity. While this implies that robots may not be able to complete their assigned tasks, it is consistent with the idea that basal rates r_{i0} are critical for task completion. Global convergence to the tasks will require a thorough characterization of the critical points of ϕ_i in (22), or even different mobility schemes, possibly optimization-based such as Model Predictive Control, as well as novel integration techniques with the communication variables.

IV. Algorithm Analysis

This section shows that Algorithm 1 ensures approximate network integrity in the sense of Problem 2 for all times t. We start in Section IV-A by analyzing the dual gradient descent portion of Algorithm 1 as summarized in steps 3 through 5. We derive an explicit bound on the suboptimality of dual iterates $\boldsymbol{\lambda}(t_n)$ in terms of problem constants [cf. Theorem 1], which we leverage to bound the infeasibility and suboptimality of primal iterates $T_{ij}(t_n)$ and $r_i(t_n)$ [cf. corollaries 1 and 2]. In Section IV-B we analyze the mobility part of Algorithm 1 as given by Step 6 to show approximate network integrity within the tolerance e in (19). We characterize e explicitly in terms of system constants [cf. Theorem 2].

A. Feasibility and Optimality of Communication Variables

To simplify presentation introduce a vector $\mathbf{t}$ stacking the rows of the transmission probability matrix $\mathbf{T}$ and a matrix $\mathbf{A}_{\mathbf{x}}$ with dimensions $J \times J(J+K)$ so as to write the constraints in (3) as $\mathbf{A}_{\mathbf{x}}\mathbf{t} - \mathbf{r} \geq \mathbf{0}$. It is not difficult to see that for this to be true it must be $A_{i,i+j}(\mathbf{x}) = R(\mathbf{x}_i, \mathbf{x}_j)$ for all $i = 1, \ldots, J$ and $j = 1, \ldots, J+K$, whereas $A_{i,(n-1)j+i}(\mathbf{x}) = -R(\mathbf{x}_j, \mathbf{x}_i)$ for all $i = 1, \ldots, J$ and $j = 1, \ldots, J+K$ with $j \neq i$. The specific form of this mapping is not important, however, as we are just using $\mathbf{A}_{\mathbf{x}}\mathbf{t} - \mathbf{r} \geq \mathbf{0}$ as shorthand notation for the constraints in (3). Using this definition we can rewrite (4) as

$$P_{\mathbf{x}} = \max f_0(\mathbf{r}, \mathbf{t}), \quad \text{s.t.} \mathbf{A}_{\mathbf{x}}\mathbf{t} - \mathbf{r} \geq \mathbf{0}, \tag{25}$$

where the constraints $r_i \geq r_{i0}$ and $\sum_{j=1}^{J+K} T_{ij} \leq 1$ were left implicit. Likewise, we use this shorthand notation to write the optimal distributed communication algorithm in (12)–(13) as

$$\mathbf{t}(t_n), \mathbf{r}(t_n) = \operatorname{argmax} \mathcal{L}_{\mathbf{x}(t_n),i}(\boldsymbol{\lambda}(t_n), \mathbf{t}, \mathbf{r}), \tag{26a}$$

$$\boldsymbol{\lambda}(t_{n+1}) = \mathbb{P}\left[\boldsymbol{\lambda}(t_n) - \epsilon\left(\mathbf{A}_{\mathbf{x}(t_n)}\mathbf{t}(t_n) - \mathbf{r}(t_n)\right)\right]. \tag{26b}$$

In a static setting, i.e., when the robot positions $\mathbf{x}(t)$ are fixed, e.g., at $\mathbf{x}(t_n)$, for all time $t \geq t_0$, it is known that the dual variables $\boldsymbol{\lambda}(t_n)$ approach the optimal multipliers $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$. As we already observed, convergence of the primal variables $\mathbf{t}(t_n)$ and $\mathbf{r}(t_n)$ to the optimal network operating point $\mathbf{t}^*_{\mathbf{x}(t_n)}, \mathbf{r}^*_{\mathbf{x}(t_n)}$ follows provided some technical conditions hold. In the dynamic setting considered here, the primal and dual variable updates in (26) bring the network closer to its optimal operating point. However when terminals move as per (24), the optimal operating point drifts away towards $\mathbf{t}^*_{\mathbf{x}(t_{n+1})}, \mathbf{r}^*_{\mathbf{x}(t_{n+1})}$. Our goal in this section is to determine the optimality of the operating point $\mathbf{t}(t_n), \mathbf{r}(t_n)$ with respect to the optimal operating point $\mathbf{t}^*_{\mathbf{x}(t_n)}, \mathbf{r}^*_{\mathbf{x}(t_n)}$ for the current team configuration $\mathbf{x}(t_n)$.

Characterizing the optimality of operating point $\mathbf{t}(t_n), \mathbf{r}(t_n)$ concerns determination of feasibility and optimality. For the operating point $\mathbf{t}(t_n)$, $\mathbf{r}(t_n)$ to be feasible for configuration $\mathbf{x}(t_n)$ we must have $\mathbf{A}_{\mathbf{x}(t_n)}\mathbf{t}(t_n) - \mathbf{r}(t_n) \geq \mathbf{0}$. This cannot be guaranteed, but we will see that the norm of the negative components of $\mathbf{A}_{\mathbf{x}(t_n)}\mathbf{t}(t_n) - \mathbf{r}(t_n)$, i.e., the norm of the projection $\mathbb{P}[\mathbf{r}(t_n) - \mathbf{A}_{\mathbf{x}(t_n)}\mathbf{t}(t_n)]$ can be made small. To determine optimality we will bound the difference between the optimal objective $P_{\mathbf{x}(t_n)}$ for the current configuration, and the yield $f_0(t_n) \triangleq f_0(\mathbf{t}(t_n), \mathbf{r}(t_n))$ associated with the operating point $\mathbf{t}(t_n)$, $\mathbf{r}(t_n)$.

Throughout the subsequent analysis we make the following assumptions on the dual functions $g_{\mathbf{x}}(\boldsymbol{\lambda})$:

(A1) The dual functions $g_{\mathbf{x}}(\boldsymbol{\lambda})$ are strongly convex for all $\mathbf{x}$ with common strong convexity parameter m,

$$g_{\mathbf{x}}(\boldsymbol{\mu}) \geq g_{\mathbf{x}}(\boldsymbol{\lambda}) + \nabla g_{\mathbf{x}}(\boldsymbol{\lambda})^T(\boldsymbol{\mu} - \boldsymbol{\lambda}) + \frac{m}{2}\|\boldsymbol{\lambda} - \boldsymbol{\mu}\|^2. \tag{27}$$

(A2) The gradients of the dual functions $g_{\mathbf{x}}(\boldsymbol{\lambda})$ are Lipschitz continuous with common Lipschitz constant M

$$\|\nabla g_{\mathbf{x}}(\boldsymbol{\lambda}) - \nabla g_{\mathbf{x}}(\boldsymbol{\mu})\| \leq M\|\boldsymbol{\lambda} - \boldsymbol{\mu}\|. \tag{28}$$

(A3) The 2-norm of the dual gradients $\nabla g_{\mathbf{x}}(\boldsymbol{\lambda})$ are uniformly bounded for all $\boldsymbol{\lambda}$ and all $\mathbf{x}$,

$$\|\nabla g_{\mathbf{x}}(\boldsymbol{\lambda})\| \leq G_{\max}. \tag{29}$$

(A4) The 1-norm of the optimal Lagrange multipliers $\boldsymbol{\lambda}^*_{\mathbf{x}}$ are uniformly bounded for all $\mathbf{x}$

$$\|\boldsymbol{\lambda}^*_{\mathbf{x}}\|_1 \leq \lambda_{\max}. \tag{30}$$

These assumptions are mild, technical, and commonly required in the analysis of gradient descent algorithms. If the dual function is twice differentiable, assumptions (A1) and (A2) impose a lower bound m and an upper bound M on the eigenvalues of the dual Hessian. For the Lipschitz constant in (A2) to exist it suffices to require the objective function $f_0(\mathbf{r}, \mathbf{t})$ to be strongly convex. If the objective function $f_0(\mathbf{r}, \mathbf{t})$ has Lipschitz gradients and the matrix $\mathbf{A}_{\mathbf{x}}\mathbf{A}_{\mathbf{x}}^T$ is full row rank, the dual function is strongly convex as required by (A1). Since dual gradients are given by the constraint violations, and the primal variables are finite, the bound $G_{\max}$ exists. The bound (30) on the optimal dual variables assumed in (A4) exists as long as the existence of strictly feasible operating points is guaranteed.

Since the iteration in (26) implements gradient descent in the dual domain. The main result derived in this section describes the distance between the current Lagrange multiplier $\boldsymbol{\lambda}(t_n)$ and the current optimal Lagrange multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$. The results of interest for the primal variables follow as corollaries of this result that we state in the following theorem.

Theorem 1: Let $\mathbf{x}(t_n)$ denote the team configuration at iteration n, $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$ the corresponding optimal dual variable and $\boldsymbol{\lambda}(t_n)$ the dual iterate obtained through iterative application of (26). Assume the step size in (26) is bounded as $\epsilon \leq 1/M$ and that the difference between reliabilities at subsequent configurations is absolutely bounded by $\delta > 0$, i.e.,

$$\left|R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))\right| \leq \delta. \tag{31}$$

If assumptions (A1), (A2), and (A4) hold, the distance between the dual iterate $\boldsymbol{\lambda}(t_n)$ and the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$ satisfies

$$\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\| \leq \beta^n\|\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_0)}\| + \sqrt{\frac{2\lambda_{\max}J}{m(1-\beta)^2}\delta}, \tag{32}$$

where the constant β is defined as $\beta \triangleq \sqrt{1/(1+m\epsilon)}$.

Proof: See Appendix A. ∎

Theorem 1 dictates that dual iterates $\boldsymbol{\lambda}(t_n)$ converge linearly to a neighborhood of the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$. The volume of this neighborhood is determined by problem-specific constants and can be made arbitrarily small by reducing δ, which can be controlled by modulating the velocity of the robots. Before elaborating on the implications of the bound in (32) let us translate the result in Theorem 1 into results regarding primal variables. The feasibility gap of $\mathbf{t}(t_n)$, $\mathbf{r}(t_n)$ is bounded in Corollary 1 and a similar bound for the optimality gap is shown in Corollary 2.

Corollary 1: With the same hypotheses and definitions of Theorem 1, the norm of the constraint violation can be bounded as

$$\|\mathbb{P}(\mathbf{r}(t_n) - \mathbf{A}_{\mathbf{x}(t_n)}\mathbf{t}(t_n))\| \leq M\beta^n\|\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_0)}\| + \sqrt{\frac{2M^2\lambda_{\max}J}{m(1-\beta)^2}\delta}. \tag{33}$$

Proof: In this proof we consider the single configuration $\mathbf{x}(t_n)$. Thus, to simplify notation we drop the configuration subscripts in the matrix $\mathbf{A}_{\mathbf{x}(t_n)} = \mathbf{A}$, the dual function $g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}) = g(\boldsymbol{\lambda})$ and the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} = \boldsymbol{\lambda}^*$.

From the assumption that gradients are Lipschitz continuous it follows that the norm of the difference between the gradient at $\boldsymbol{\lambda}(t_n)$ and the gradient at the optimal multiplier $\boldsymbol{\lambda}^*$ is bounded as

$$\|\nabla g(\boldsymbol{\lambda}(t_n)) - \nabla g(\boldsymbol{\lambda}^*)\| \leq M\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*\| \tag{34}$$

Consider now the projections $\mathbb{P}[-\nabla g(\boldsymbol{\lambda}(t_n))]$ and $\mathbb{P}[-\nabla g(\boldsymbol{\lambda}^*)]$ of the gradients on the nonnegative orthant. Observe that because the components of the optimal dual gradient are either null or positive it holds $\mathbb{P}[-\nabla g(\boldsymbol{\lambda}^*)] = \mathbf{0}$. Further notice that since the nonnegative orthant is a convex set the distance between the projections is smaller than the distance between the vectors. Combining these two observations we can write

$$\begin{aligned}\|\mathbb{P}[-\nabla g(\boldsymbol{\lambda}(t_n))]\| &= \|\mathbb{P}[-\nabla g(\boldsymbol{\lambda}(t_n))] - \mathbb{P}[-\nabla g(\boldsymbol{\lambda}^*)]\| \\ &\leq \|\nabla g(\boldsymbol{\lambda}(t_n)) - \nabla g(\boldsymbol{\lambda}^*)\| \end{aligned} \quad (35)$$

But the gradient of the dual function *is* the constraint violation, i.e., $\nabla g(\boldsymbol{\lambda}(t_n))) = \mathbf{A}\mathbf{t}(t_n) - \mathbf{r}(t_n)$. Substituting this equality in (35) and combining with the bound in (34) yields

$$\|\mathbb{P}[-(\mathbf{A}\mathbf{t}(t_n) - \mathbf{r}(t_n))]\| \leq M\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*\|. \quad (36)$$

The result in (33) follows from substituting bound (32) of Theorem 1 for the right hand side of (36). ∎

Corollary 2: Let $f_0(t_n) \triangleq f_0(\mathbf{t}(t_n), \mathbf{r}(t_n))$ denote the primal objective corresponding to the communication variables at time t_n. With the same hypotheses and definitions of Theorem 1, the optimality gap can be bounded as

$$\begin{aligned}P_{\mathbf{x}(t_n)} - f_0(t_n) \leq (G_{\max} + \lambda_{\max}M)^2\beta^n\|\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_0)}\| \\ + \sqrt{\frac{2(G_{\max} + \lambda_{\max}M)\lambda_{\max}J}{m(1-\beta)^2}\delta}. \end{aligned} \quad (37)$$

Proof: As in the proof of Corollary 1 simplify notation to $\mathbf{A}_{\mathbf{x}(t_n)} = \mathbf{A}$, $g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}) = g(\boldsymbol{\lambda})$ and $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} = \boldsymbol{\lambda}^*$. We further adopt the convention $P = P_{\mathbf{x}(t_n)}$. Since variables $\mathbf{t}(t_n)$ and $\mathbf{r}(t_n)$ are Lagrangian maximizers for multiplier $\boldsymbol{\lambda}(t_n)$, the dual function evaluated at $\boldsymbol{\lambda}(t_n)$ can be written as

$$g(\boldsymbol{\lambda}(t_n)) = f_0(\mathbf{t}(t_n), \mathbf{r}(t_n)) + \boldsymbol{\lambda}(t_n)^T(\mathbf{A}\mathbf{t}(t_n) - \mathbf{r}(t_n)) \quad (38)$$

Likewise, the dual function evaluated at $\boldsymbol{\lambda}^*$ is given by

$$g(\boldsymbol{\lambda}^*) = P + \boldsymbol{\lambda}^{*T}(\mathbf{A}\mathbf{t}^* - \mathbf{r}^*) \quad (39)$$

Subtract (38) from (39), use the fact that $g(\boldsymbol{\lambda}(t_n)) - g(\boldsymbol{\lambda}^*) \geq 0$, recall that constraint violations are dual function gradients, i.e., $\mathbf{A}\mathbf{t}(t_n) - \mathbf{r}(t_n) = \nabla g(\boldsymbol{\lambda}(t_n))$ and $\mathbf{A}\mathbf{t}^* - \mathbf{r}^* = \nabla g(\boldsymbol{\lambda}^*)$, and substitute the definition $f_0(t_n) \triangleq f_0(\mathbf{t}(t_n), \mathbf{r}(t_n))$ to obtain the bound

$$P_{\mathbf{x}(t_n)} - f_0(t_n) \leq \boldsymbol{\lambda}(t_n)^T\nabla g(\boldsymbol{\lambda}(t_n)) - \boldsymbol{\lambda}^{*T}\nabla g(\boldsymbol{\lambda}^*) \quad (40)$$

Add and subtract $\boldsymbol{\lambda}^{*T}\nabla g(\boldsymbol{\lambda}(t_n))$ from the right hand side of (40) and reorder terms to write

$$\begin{aligned}P_{\mathbf{x}(t_n)} - f_0(t_n) \leq \left(\boldsymbol{\lambda}(t_n)^T - \boldsymbol{\lambda}^*\right)\nabla g(\boldsymbol{\lambda}(t_n)) \\ + \boldsymbol{\lambda}^{*T}(\nabla g(\boldsymbol{\lambda}(t_n)) - \nabla g(\boldsymbol{\lambda}^*)) \end{aligned} \quad (41)$$

Using the assumption that the norm of dual gradients is bounded by $G_{\max}$, the first summand in the right hand side of (41) can be bounded as

$$\left(\boldsymbol{\lambda}(t_n)^T - \boldsymbol{\lambda}^*\right)\nabla g(\boldsymbol{\lambda}(t_n)) \leq G_{\max}\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*\| \quad (42)$$

Recall now the assumption that the norm of optimal dual gradients are bounded by $\lambda_{\max}$ (the 2-norm of a vector is smaller than its 1-norm). Further use the assumption that dual gradients are Lipschitz to write

$$\boldsymbol{\lambda}^{*T}(\nabla g(\boldsymbol{\lambda}(t_n)) - \nabla g(\boldsymbol{\lambda}^*)) \leq \lambda_{\max}M\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*\| \quad (43)$$

Substitute now the bounds in (42) and (43) into the inequality in (41) to obtain

$$P_{\mathbf{x}(t_n)} - f_0(t_n) \leq (G_{\max} + \lambda_{\max}M)\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*\|. \quad (44)$$

The result in (37) follows from substituting bound (32) of Theorem 1 for the right hand side of (44). ∎

The results in Theorem 1 and Corollaries 1 and 2 are conceptually similar. Both of them include a term characterizing the vanishment of the initial optimality gap $\|\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_0)}\|$ and the steady state gap of the achieved operating point $\boldsymbol{\lambda}(t_n)$, $\mathbf{t}(t_n)$, $\mathbf{r}(t_n)$. This steady state gap defines a neighborhood of the (time-varying) optimal operating point to which the iterates defined by (26) converge.

Convergence towards this neighborhood is determined by the constant β, which is the same constant that would characterize convergence in a static setting. The stepsize can range in the interval $(0, 1/M]$ resulting in constants β^2 ranging in the interval $[1/(1 + m/M), 1)$. Naturally, smaller stepsizes result in constants β closer to 1, thereby decreasing the convergence rate. Smaller stepsizes also decrease the steady state gap between the optimal and achieved network operating points. The most convenient stepsize selection is therefore to make $\epsilon = 1/M$. In such case the constant β^2 is as small as it can be and is determined by the ratio $M/m \geq 1$. This ratio is the condition number of an optimization problem and describes how close to a hypersphere the level sets of the dual function are. The most favorable case is when $m = M$, corresponding to a dual function having hyperspheres as level sets, and yielding $\beta^2 = 1/2$. For problems with smaller condition numbers the convergence rate decreases and the steady state performance gap increases. This condition number is determined by the selection of the primal objective.

The steady state performance gap also increases with the size of the network J and the bound $\lambda_{\max}$ on the norm of the optimal gradients. This latter dependence is interesting because the norm of the optimal multiplier is related to the difficulty in satisfying problem constraints. Constraints that are difficult to satisfy are characterized by a large sensitivity to parameter perturbations. This sensitivity is determined by the value of the corresponding optimal Lagrange multiplier.

B. Network Integrity

Theorem 1 and corollaries 1 and 2 guarantee that the communication variables are close to optimality and feasibility at all sampling times t_n. This proximity depends on the magnitude of the disturbance $\left|R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))\right| = \delta > 0$ that is introduced in (4) due to robot mobility between consecutive communication updates at times t_n and t_{n+1}. In this section we extend the feasibility guarantees of Theorem 1 to all times $t \geq t_0$. For this, assume that the channel reliabilities $R(\mathbf{x}_i(t), \mathbf{x}_j(t))$ are purely functions of the inter-robot distances $\|\mathbf{x}_i(t) - \mathbf{x}_j(t)\|$ and let

$$\max_{i,j}\left\{\left|\frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\right|\right\} \leq M_R < \infty, \quad (45)$$

where $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$, denote an upper bound on their absolute maximum slope. Let also

$$\max_i \{\|\nabla_{\mathbf{x}_i}\phi_i\|\} \leq M_\phi < \infty \tag{46}$$

denote an upper bound on the maximum robot speed; see (24). It is not difficult to show that such a bound M_ϕ for the proposed potential exists; see, e.g., [36], [37]. Finally, define the piecewise constant functions of time $T_{ij}(t) = T_{ij}(t_n)$, $\lambda_i(t) = \lambda_i(t_n)$ and $\sigma_i(t) = \sigma_i(t_n)$ for all times $t \in [t_n, t_{n+1})$ and all indices i and j, and recall that $\mathcal{F}(\sigma(t))$ with $\sigma(t) = \{\sigma_1(t), \ldots, \sigma_J(t)\}$ denotes the set of robot configurations $\mathbf{x}(t)$ that ensure communication rates $\beta_i(\mathbf{x}(t); \sigma_i(t))$ that exceed desired basal rates r_{i0} with some error $e > 0$, i.e., $\beta_i(\mathbf{x}(t); \sigma_i(t)) > r_{i0} - e$ [cf. (20)]. Then, we have the result:

Theorem 2: Define $\Delta t \triangleq \max_n\{t_{n+1} - t_n\}$ and let M_R be the bound on the norm of the gradient of $R(\mathbf{x}_i, \mathbf{x}_j)$ in (45) and M_ϕ the bound on robots' velocities in (46). Assume the hypotheses in Theorem 1 hold and that for time t_0 the communication variables are initialized at the optimal configuration, i.e., $\boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}^*$, $\mathbf{r}(t_0) = \mathbf{r}^*_{\mathbf{x}(t_0)}$, and $\mathbf{T}(t_0) = \mathbf{T}^*_{\mathbf{x}(t_0)}$. Then, for any tolerance e satisfying

$$e > \sqrt{2\alpha\Delta t M_R M_\phi \frac{2M^2\lambda_{\max}}{m(1-\beta)^2}}, \tag{47}$$

we have $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all times $t \geq t_0$.

Proof: The proof of this result consists of two parts. We commence bounding the disturbance $\big|R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))\big| = \delta > 0$ in terms of the problem specific constants M_R and M_ϕ. This allows us to explicitly characterize the approximation trade-offs that arise due to robot mobility. We then show that these bounded feasibility errors can be guaranteed for all time $t \geq t_0$. I.e., we show that the set $\mathcal{F}(\sigma(t))$ is an invariant of motion for the closed loop system.

Let us begin by deriving a bound for the disturbance $\delta > 0$ in terms of the problem specific parameters M_R and M_ϕ. Observe that for all times $t \geq t_0$ we have

$$\dot{R}(\mathbf{x}_i, \mathbf{x}_j) = \frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\frac{d}{dt}\|\mathbf{x}_{ij}\| = \frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\frac{\mathbf{x}_{ij}^T\dot{\mathbf{x}}_{ij}}{\|\mathbf{x}_{ij}\|}$$

and substituting $\dot{\mathbf{x}}_i = -\alpha\nabla_{\mathbf{x}_i}\phi_i$ from (24), we get

$$\dot{R}(\mathbf{x}_i, \mathbf{x}_j) = -\alpha\frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\frac{\mathbf{x}_{ij}^T(\nabla_{\mathbf{x}_i}\phi_i - \nabla_{\mathbf{x}_j}\phi_j)}{\|\mathbf{x}_{ij}\|}. \tag{48}$$

Taking the absolute value of both sides of (48) and applying concecutively the Cauchy-Schwarz inequality and the triangle inequality, we obtain

$$\begin{aligned}|\dot{R}(\mathbf{x}_i, \mathbf{x}_j)| &\leq \alpha\left|\frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\right|\frac{\|\mathbf{x}_{ij}\|\left\|\nabla_{\mathbf{x}_i}\phi_i - \nabla_{\mathbf{x}_j}\phi_j\right\|}{\|\mathbf{x}_{ij}\|}\\ &\leq \alpha\left|\frac{dR(\mathbf{x}_i, \mathbf{x}_j)}{d\|\mathbf{x}_{ij}\|}\right|\left(\|\nabla_{\mathbf{x}_i}\phi_i\| + \left\|\nabla_{\mathbf{x}_j}\phi_j\right\|\right)\\ &\leq 2\alpha M_R M_\phi,\end{aligned} \tag{49}$$

where the last step follows from (45) and (46). Expressing the disturbance $R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))$ as a definite integral, we have that

$$\begin{aligned}&\big|R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))\big| =\\ &= \Big|\int_{t_n}^{t_{n+1}} \dot{R}(\mathbf{x}_i(\tau), \mathbf{x}_j(\tau))d\tau\Big|\\ &\leq \int_{t_n}^{t_{n+1}} |\dot{R}(\mathbf{x}_i(\tau), \mathbf{x}_j(\tau))|d\tau\\ &\leq \int_{t_n}^{t_{n+1}} 2\alpha M_R M_\phi d\tau = 2\alpha(t_{n+1} - t_n)M_R M_\phi,\end{aligned} \tag{50}$$

with the second inequality due to (49). Observe now that since $\Delta t = \max_n\{t_{n+1} - t_n\}$, (50) we an write

$$\big|R(\mathbf{x}_i(t_{n+1}), \mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n), \mathbf{x}_j(t_n))\big| \leq 2\alpha\Delta t M_R M_\phi \tag{51}$$

from which we obtain that the bound $\delta = 2\alpha\Delta t M_R M_\phi$ in Theorem 1. This concludes the first part of the proof.

For the second part of the proof, we decompose time into the sampling times $\{t_n\}_{n=0}^\infty$ and the collection of motion intervals $[t_0, \infty)\backslash\{t_n\}_{n=0}^\infty$. Based on Corollary 1 and (51), we first determine a bound on the error $e > 0$ that ensures that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in \{t_n\}_{n=0}^\infty$. Then, given this bound and the fact that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in \{t_n\}_{n=0}^\infty$, we show that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in [t_0, \infty)\backslash\{t_n\}_{n=0}^\infty$. Combining the two results, we get that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all time $t \in [t_0, \infty)$, which proves our claim.

Let us begin by assuming that $t \in \{t_n\}_{n=0}^\infty$ and recall the hypothesis that for time t_0 the communication variables are initialized at the optimal configuration, i.e., $\boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}^*$, $\mathbf{r}(t_0) = \mathbf{r}^*_{\mathbf{x}(t_0)}$, and $\mathbf{T}(t_0) = \mathbf{T}^*_{\mathbf{x}(t_0)}$. This means that the first term on the right-hand-side of (32) vanishes for all $n > 0$. Then, (33) in Corollary 1 can be equivalently written

$$\sqrt{\sum_{i=1}^J \mathbb{P}\left[r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t))\right]^2} \leq \sqrt{2\alpha\Delta t M_R M_\phi \frac{2M^2\lambda_{\max}J}{m(1-\beta)^2}}, \tag{52}$$

where $\beta_i(\mathbf{x}(t); \sigma_i(t))$ is defined in (16) and $\delta = 2\alpha\Delta t M_R M_\phi$, and $\mathbb{P}[\cdot]$ is the projection on the non-negative orthant. A sufficient condition for (52) to hold is that

$$\mathbb{P}\left[r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t))\right]^2 \leq 2\alpha\Delta t M_R M_\phi \frac{2M^2\lambda_{\max}}{m(1-\beta)^2}. \tag{53}$$

for all $i = 1, \ldots, J$. Taking the square root of both sides of (53) and using the fact that $\mathbb{P}\left[r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t))\right] = \max\{0, r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t))\} \geq 0$ we get

$$r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t)) \leq \sqrt{2\alpha\Delta t M_R M_\phi \frac{2M^2\lambda_{\max}}{m(1-\beta)^2}}. \tag{54}$$

Clearly, if $r_i(t) - \beta_i(\mathbf{x}(t); \sigma_i(t))$ is upper bounded as in (54), then it is also upper bounded by any $e > 0$ that satisfies (47).

The above argument also shows that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in \{t_n\}_{n=0}^\infty$. To see this, note that for $e > 0$ defined as in (47), the bound in (54) implies that $\beta_i(\mathbf{x}(t); \sigma_i(t)) - r_i(t) + e > 0$

for all $t \in \{t_n\}_{n=0}^{\infty}$. Since $r_i(t) \geq r_{i0}$ and $\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t)) = \beta_i(\mathbf{x}(t);\sigma_i(t)) - r_{i0} + e$ by (19), we obtain that $\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t)) > 0$ for all $t \in \{t_n\}_{n=0}^{\infty}$. Therefore, $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in \{t_n\}_{n=0}^{\infty}$, as desired.

We next show that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in [t_0,\infty)\backslash\{t_n\}_{n=0}^{\infty}$. For this to be true, every time interval (t_n, t_{n+1}) should be initialized with the robot positions in the free space, i.e., with $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all $t \in \{t_n\}_{n=0}^{\infty}$. As shown above, this initialization can be guaranteed if the error $e > 0$ is chosen as in (47). Assume now that $\mathbf{x}(t) \in \mathbb{R}^{Jd}$ is such that $\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t)) = 0$ for any $t \in [t_0,\infty)\backslash\{t_n\}_{n=0}^{\infty}$ and any $i = 1,\ldots,J$. Then, for leader robot i we have

$$\phi_i(\mathbf{x}(t);\sigma_i(t)) = \frac{\gamma_i(\mathbf{x}_i(t))}{(\gamma_i^k(\mathbf{x}_i(t)) + \tilde{\beta}_i^2(\mathbf{x}(t);\sigma_i(t)))^{1/k}} = 1,$$

which means that at $\mathbf{x}(t)$ for $t \in [t_0,\infty)\backslash\{t_n\}_{n=0}^{\infty}$, the potential ϕ_i achieves its maximum. Since the set of initial conditions $\mathcal{F}(\sigma(t))$ is open for any $t \in \{t_n\}_{n=0}^{\infty}$, and no open set of initial conditions can be attracted to the maxima of ϕ_i along the negative gradient motion $-\nabla_{\mathbf{x}_i}\phi_i$ defined by the closed loop system (24), we also conclude that $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for $t \in [t_0,\infty)\backslash\{t_n\}_{n=0}^{\infty}$. Therefore, $\mathbf{x}(t) \in \mathcal{F}(\sigma(t))$ for all time $t \geq t_0$. The same argument holds for relay robots, by setting $\gamma_i = 1$. ■

The error lower bound in Theorem 2 can be made arbitrarily small by either reducing the robot speed gain $\alpha > 0$ or decreasing the maximum interval $\Delta t \geq t_{n+1} - t_n$ between updates of routing variables. Reducing this error bound reduces the violation of the network integrity constraints, thereby bringing achieved communication rates closer to their minimum requirements r_{i0}. Equivalently, for a given error tolerance e and frequency $1/\Delta t$ of communication variable updates, the speed gain $\alpha > 0$ needs to be sufficiently small to realize the desired bounds in (47). In that situation the choice of the lower bound in Theorem 2 introduces a tradeoff between optimal communication and robot speed. Theorem 2 expresses this tradeoff in terms of problem parameters.

V. Numerical Simulations

We illustrate the proposed distributed controller as summarized in Algorithm 1 with computer simulations where communication integrity of the robot network needs to be preserved. We employ channel reliabilities $R(\mathbf{x}_i,\mathbf{x}_j)$ that are deterministic functions of the distances between robots $\|\mathbf{x}_i - \mathbf{x}_j\|$ and satisfy $R(\mathbf{x}_i,\mathbf{x}_j) = 1$ if $\|\mathbf{x}_{ij}\| \leq l$,

$$R(\mathbf{x}_i,\mathbf{x}_j) = a\|\mathbf{x}_{ij}\|^3 + b\|\mathbf{x}_{ij}\|^2 + c\|\mathbf{x}_{ij}\| + d, \qquad (55)$$

if $l < \|\mathbf{x}_{ij}\| \leq u$ and $R(\mathbf{x}_i,\mathbf{x}_j) = 0$ if $\|\mathbf{x}_{ij}\| > u$, where $\mathbf{x}_{ij} \triangleq \mathbf{x}_i - \mathbf{x}_j$. The constants $0 < l < u$ are lower and upper bounds on the robot distances and the constants a, b, c, and d are chosen to make $R(\mathbf{x}_i,\mathbf{x}_j)$ a twice differentiable function ranging from 0 to 1; see Fig. 3. This function is a polynomial fitting of experimental curves found in the literature, e.g., [38].

Consider a scenario where reliable communications is to be established between a single AP located at the origin $(0,0) \in \mathbb{R}^2$ and a service point (SP) located at $(0.44,0) \in \mathbb{R}^2$. This task is performed by a team of $J = 6$ point robots, randomly initialized according to a uniform distribution in a square of side 0.04 units, centered 0.06 units to the right of the AP; see Fig. 5(a). This makes the initial position of the robots close enough to the AP to ensure the communication network is initially connected. Robot 2 is designated as the leader and tasked with serving the SP. All robots $i = 2,\ldots,7$ relay information back to the AP. The AP is indexed as $i = 1$. Robot motion is according to the closed loop dynamics in (24). The task potential γ_i in (22) for the leader robot $i = 2$ is defined as $\gamma_i = \|\mathbf{x}_i - \mathbf{x}_i^d\|_2^2$ with $\mathbf{x}_i^d = (0.44, 0)$ corresponding to the coordinates of the SP. The switching signal $\sigma_i(t_n)$ in (15) is updated through the primal-dual iteration in (12)–(13). This is equivalent to the summary in Algorithm 1. The utility functions of problem (4) are taken to be $U_i(r_i) = \log(r_i)$ and $V_{ij}(T_{ij}) = -T_{ij}^2$, so that the resulting sum utility $\sum_{i=1}^{J} U_i(r_i) + \sum_{j=1}^{J+K} V_{ij}(T_{ij})$ is strictly concave, as required by Assumption (A1) in Section IV.

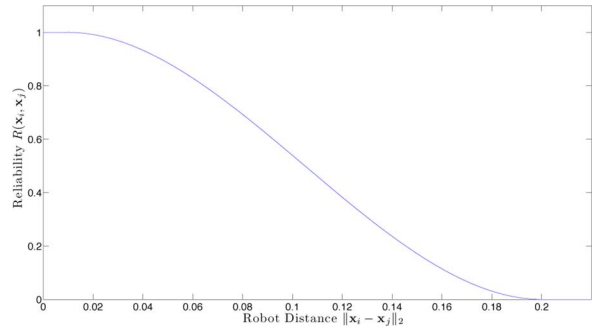


Fig. 3. Channel reliability function $R(\mathbf{x}_i,\mathbf{x}_j)$ used in the simulations in Section V. Reliabilities are assumed deterministic functions of distances between robots $\|\mathbf{x}_i - \mathbf{x}_j\|$ as in (55). Here $l = 0.01$ and $u = 0.2$.

The remaining simulation parameters are defined as follows: (a) The minimum rates r_{i0} are identically zero for all robots except for the leader for which $r_{i0} = 0.7$. This choice is consistent with the classification of robots into relays and leaders. Leaders collect measurements and generate data, while relay robots convey this information back to the APs. (b) The error tolerance $e > 0$ in (18) is set to $e = 0.1$. Note that estimating e directly from (47) is not easy, since explicit values for many problem specific parameters are difficult to obtain. Moreover, the provided lower bound on e is rather conservative and its value should be interpreted in an existential way. (c) The constant $k > 0$ in (22) is set to $k = 80$. This value is large enough so that the effect of the obstacle potentials $\tilde{\beta}_i(\mathbf{x}(t);\sigma_i(t_n))$ is restricted to a small area around the boundary $\partial\mathcal{F}(\sigma(t))$ of the free space $\mathcal{F}(\sigma(t))$, as defined in (21); see also [36], [37]. This allows the leader robot to move freely towards its assigned SP unless communication constraints are close to being violated. (d) The integration step size in Theorem 2 is $\Delta t = 10^{-4}$ and the controller gain in (23) is set to $\alpha = 10^2$.

We further add potentials to avoid collisions between nearby robots. Collision potentials $\zeta_{ij} : \mathbb{R}^{Jd} \to \mathbb{R}_+$ are defined as

$$\zeta_{ij}(\mathbf{x}_i,\mathbf{x}_j) = \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\eta + \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \qquad (56)$$

with $\eta > 0$; see Fig. 4. The effect of the collision potentials ζ_{ij} is more localized for smaller $\eta > 0$. We use $\eta = 10^{-7}$ for subsequent simulations. Incorporating the collision potentials, the overall artificial potential ϕ_i in (22) for motion control becomes

$$\phi_i = \begin{cases} \frac{\gamma_i}{(\gamma_i^k + G_i)^{1/k}}, & \text{if } i \text{ is a leader,} \\ \frac{1}{(1+G_i)^{1/k}}, & \text{if } i \text{ is a relay,} \end{cases} \qquad (57)$$

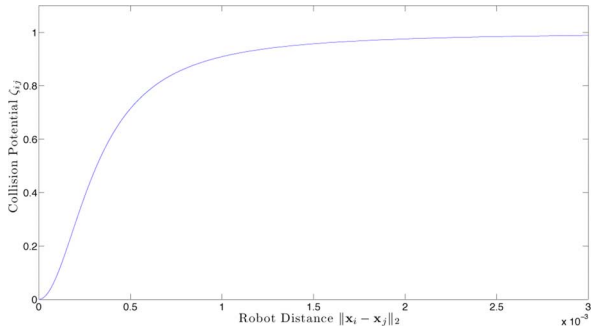


Fig. 4. Collision potential $\zeta_{ij}(\mathbf{x}_i, \mathbf{x}_j)$ used for simulations in Section V. The potential $\zeta_{ij}(\mathbf{x}_i, \mathbf{x}_j)$ is given in (56). Notice how the effect of the collision potential is negligible when the robot distance $\|\mathbf{x}_i - \mathbf{x}_j\|$ becomes larger than $3 \cdot 10^{-3}$ as the gradient becomes $\nabla_{\mathbf{x}_i} \zeta_{ij}(\mathbf{x}_i, \mathbf{x}_j) \approx 0$ ($\eta = 10^{-7}$).

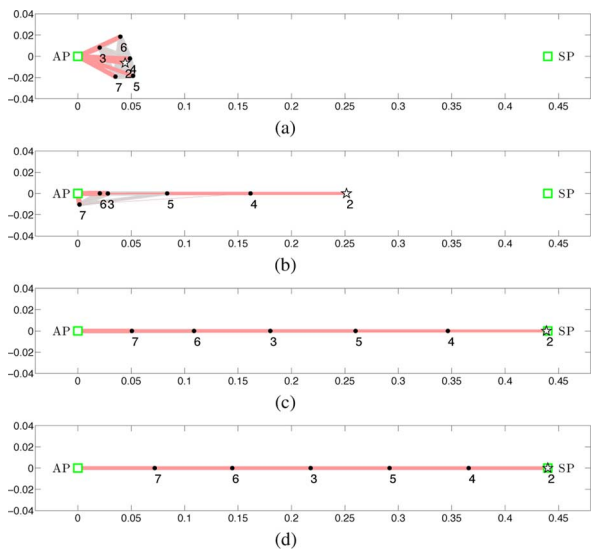


Fig. 5. A mobile robot network consisting of 6 robots (dots) that establish reliable communications between one service point (SP) and one access point (AP). Red lines represent routing of information $T_{ij} R(\mathbf{x}_i, \mathbf{x}_j)$, between pairs of robots. The thickness of each line represents the magnitude of these quantities. The network is randomly initialized close to the AP (Fig. 5(a)) and involves one leader (star), labeled as 2, that moves towards the SP (a) Time $t = 0$ (b) Time $t = 0.5$ (c) Time $t = 2.5$ (d) Time $t = 5$.

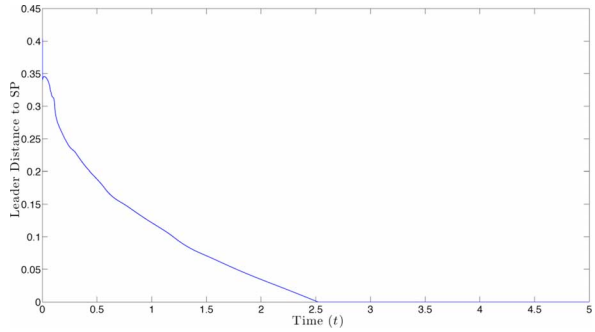


Fig. 6. Distance $\|\mathbf{x}_i - \mathbf{x}_i^d\|$ between the leader robot i and its assigned SP $\mathbf{x}_i^d = (0.44, 0)$ as a function of time t, for the scenario considered in Fig. 5.

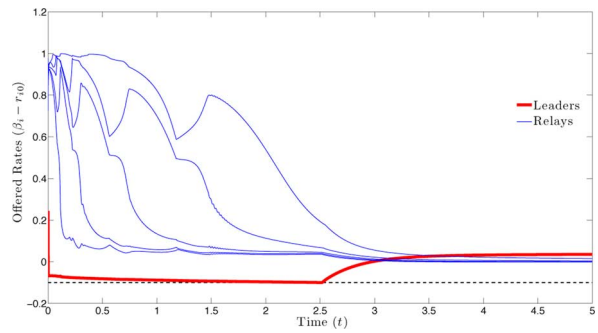


Fig. 7. Difference $\beta_i - r_{i0}$ between the offered β_i and basal r_{i0} rates for all robots i involved in the scenario in Fig. 5.

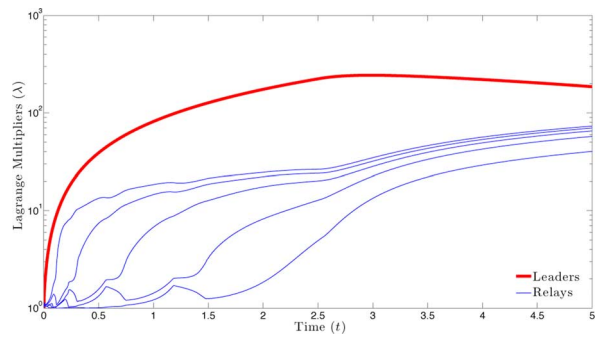


Fig. 8. Lagrange multipliers λ_i associated with the rate constraints in (3) for all robots i involved in the scenario in Fig. 5.

with $G_i = \tilde{\beta}_i^2 \prod_{j=1}^{J+K} \zeta_{ij}$. With this addition, the free space becomes the set of state variables $\mathbf{x}(t)$ and $\sigma(t)$ such that $G_i(t) > 0$ for all robots i. Note that incorporating collision avoidance does not affect the proof of Theorem 2. To see why this is true observe that $\phi_i = \gamma_i / (\gamma_i^k + G_i)^{1/k}$ becomes uniformly maximal (equal to 1) whenever either $\tilde{\beta}_i \to 0$ or $\zeta_{ij} \to 0$ for some robot $j \neq i$. Therefore, the same argument used in Theorem 2 to show communication integrity, i.e., $\tilde{\beta}_i > 0$, can be applied here to show collision avoidance, i.e., $\zeta_{ij} > 0$.

The evolution of the robot network reacting to the proposed control law summarized in Algorithm 1 is shown in the snapshots in Fig. 5 for selected time instances. Fig. 6 through 9 illustrate several metrics that evaluate its performance that we discuss in detail in Section V-A through V-D.

A. Task Completion

Even though task completion is a secondary objective, in that we provide guarantees on network integrity but not on task completion, the leader robot is still able to reach the service point (SP); see Fig. 5. A communication path is established and maintained throughout the process. Task completion can also be verified from Fig. 6 that plots the distance $\|\mathbf{x}_i - \mathbf{x}_i^d\|$ between the leader robot i and its assigned SP at $\mathbf{x}_i^d = (0.44, 0)$ as a function of time t. For the scenario under consideration, the leader reaches its destination at time $t \approx 2.5$. After this time the relay robots keep reconfiguring to improve the communication rate.

B. Performance of Communications

The performance of Algorithm 1 in terms of its ability to maintain reliable communications is illustrated in Figs. 7 and 8. Fig. 7 plots the difference $\beta_i - r_{i0}$ between the offered rates $\beta_i = \sum_{j=1}^{J+K} T_{ij} R(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^{J} T_{ji} R(\mathbf{x}_j, \mathbf{x}_i)$ in (16) and the basal rates r_{i0}, for all robots i. The horizontal dashed line corresponds to the threshold for the allowable error $e = 0.1$. Notice that the design requirement to have $\beta_i - r_{i0} > -e$ at all times is satisfied. The Lagrange multipliers λ_i associated with the rate

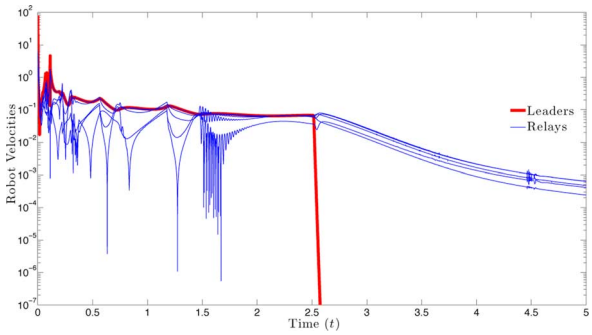


Fig. 9. Velocities $\|\dot{\mathbf{x}}_i\|$ as a function of time t, for the scenario in Fig. 5.

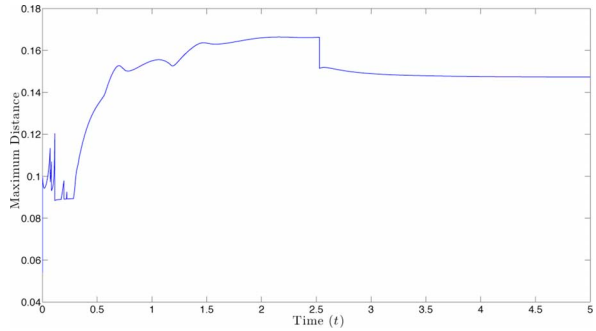


Fig. 10. Max inter-robot distance $\max_{ij}\{\|\mathbf{x}_i-\mathbf{x}_j\| | T_{ij}R(\mathbf{x}_i,\mathbf{x}_j)>0\}$.

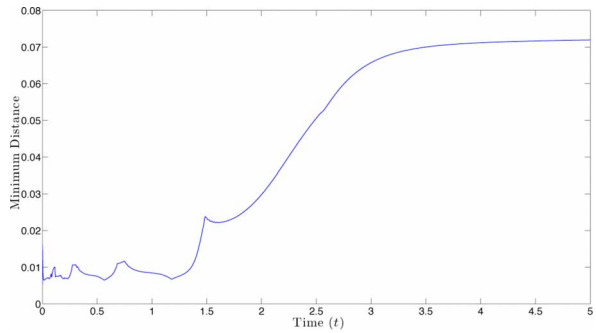


Fig. 11. Min inter-robot distance $\min_{ij}\{\|\mathbf{x}_i-\mathbf{x}_j\| | T_{ij}R(\mathbf{x}_i,\mathbf{x}_j)>0\}$.

constraints $r_i \leq \sum_{j=1}^{J+K} T_{ij}R(\mathbf{x}_i,\mathbf{x}_j) - \sum_{j=1}^{J} T_{ji}R(\mathbf{x}_j,\mathbf{x}_i)$ in (3) are shown in Fig. 8. Comparing with Fig. 7, one observes that as the constraints tend to be violated, i.e., as $\beta_i - r_{i0} \to 0$, the associated Lagrange multipliers increase. This is more apparent for the leader for which initially $\beta_i - r_{i0} < 0$. When the leader has reached its destination at $t \approx 2.5$, its Lagrange multiplier begins to decrease and the associated rate $\beta_i - r_{i0}$ increases, eventually approaching zero.

C. Effect of Robot Motion

In this section, we discuss how mobility affects communications. Equivalently, we study how much the robots can move while satisfying the communication constraints with a bounded error. Mathematically, this is captured in Theorem 1 via the upper bound $\delta > 0$ on the rate of change of the reliabilities $\big|R(\mathbf{x}_i(t_{n+1}),\mathbf{x}_j(t_{n+1})) - R(\mathbf{x}_i(t_n),\mathbf{x}_j(t_n))\big|$. However, since the explicit bound $\delta = 2\alpha\Delta t M_R M_\phi$ derived in Theorem 2 is only a conservative estimate, we instead monitor the robot velocities $\|\dot{\mathbf{x}}_i\|$ as a measure of robot motion. These velocities are plotted in Fig. 9 for all robots as a function of time. Note that the closed loop system $\dot{\mathbf{x}}_i = -\alpha\nabla_{\mathbf{x}_i}\phi_i$ in (24) implies that $\|\dot{\mathbf{x}}_i\| = \alpha\|\nabla_{\mathbf{x}_i}\phi_i\|$ and, therefore, $\|\dot{\mathbf{x}}_i\| \leq \alpha M_\phi$ by (46). This means that the robot velocities $\|\dot{\mathbf{x}}_i\|$ are proportional to the bound δ and, hence, they can be used as a proxy for δ. From Fig. 9 one observes an initial peak of the leader velocity at until about $t = 10^{-3}$ that is due to: (i) Its initially large distance from the SP. (ii) The absence of significant attractive forces from the network integrity constraints, which are easily satisfied during that time; see Figs. 7 and 8. (iii) The presence of repulsive forces from other robots in its proximity. The leader's velocity becomes 0 at about $t \approx 2.5$ when it reaches its destination, as also seen in Fig. 6. However, the remaining robots continue to move slowly to increase the offered rates β_i away from r_{i0}, see Fig. 7, and decrease the associated Lagrange multipliers, see Fig. 8. The average robot velocity throughout the exercise is between 0.1 and 0.5. With this robot speed, Algorithm 1 maintains $\beta_i - r_{i0} > -e$; see Fig. 7.

An additional observation from Fig. 9 is the rapid decrease in the leader's velocity once it reaches its destination; see also the nonsmooth change in the distance to the SP at $t \approx 2.5$ in Fig. 6. This is the result of $\gamma_i \to 0$ and $\tilde{\beta}_i \to 0$ simultaneously for the leader robot, as it can be also verified from Fig. 7. Notice that the simultaneous vanishing of γ_i and $\tilde{\beta}_i$ is not allowed in the proposed construction, see, e.g., [36], [37], therefore, the network integrity guarantees in Theorem 2 do not apply at this point. If one wishes to avoid this scenario, simple modifications are possible. One way to avoid $\gamma_i \to 0$ and $\tilde{\beta}_i \to 0$ simultaneously is to decrease the leader's speed by, e.g., decreasing the controller gain α in (23). This would give Algorithm 1 enough time to reposition the robots and reroute information to ensure that $\tilde{\beta}_i > 0$ at all times. This approach can be impractical, however, because it may lead to very slow network evolution. Another solution to this problem is to redefine the task potential as $\gamma_i = \left(\|\mathbf{x}_i - \mathbf{x}_i^d\| + b\right)^2$ for some small positive constant $b > 0$. This would prevent $\gamma_i \to 0$ and $\tilde{\beta}_i \to 0$ simultaneously but would also prevent the leader from exactly reaching the SP.

D. Geometric Considerations

Fig. 10 shows the maximum length of active links in the network. Formally, active links are ordered pairs of robots $i \to j$ for which $T_{ij}R(\mathbf{x}_i,\mathbf{x}_j) > 0$. The maximum length of all active links in the network is then defined by

$$\mathcal{A}_{\max} \triangleq \max_{ij} \{\|\mathbf{x}_i - \mathbf{x}_j\| \mid T_{ij}R(\mathbf{x}_i,\mathbf{x}_j) > 0\}.$$

Similarly, the minimum length of an active link is defined as

$$\mathcal{A}_{\min} \triangleq \min_{ij} \{\|\mathbf{x}_i - \mathbf{x}_j\| \mid T_{ij}R(\mathbf{x}_i,\mathbf{x}_j) > 0\}.$$

This minimum distance is shown in Fig. 11.

The plots in Figs. 10 and 11 can be used to extract information regarding the performance of Algorithm 1 in terms of both communication and motion control. In particular, notice that in Fig. 10 the maximum inter-robot distance converges at about $\|\mathbf{x}_i - \mathbf{x}_j\| \approx 0.14$, where $R(\mathbf{x}_i,\mathbf{x}_j) \approx 0.23$. Comparing with Fig. 5(d), we see that the second order neighbors, namely those neighbors two hops away as plotted in Fig. 5(d),

are still in communication range. However, the thickness of the lines with the first order neighbors indicates that all information is routed to them. This result is also observed in Fig. 11. Notice that the minimum inter-robot distance converges at about $\|\mathbf{x}_i - \mathbf{x}_j\| \approx 0.072$, where $R(\mathbf{x}_i, \mathbf{x}_j) \approx 0.75 \gg 0.23$. In fact, since for the leader $r_{i0} = 0.7$ and this information needs to be routed to the AP via the path shown in Fig. 5(d), we conclude that at equilibrium $\beta_i - r_{i0} \leq 0.05$ for all robots. This also explains the value of the plots in Fig. 7 as the time approaches $t = 5$. Also notice that the minimum inter-robot distance observed at the early stages of the simulation in Fig. 11, i.e., up until time $t \approx 1.5$, is about $\|\mathbf{x}_i - \mathbf{x}_j\| \approx 0.006$. Comparing with the plot of the collision potentials in Fig. 4, we see that this distance is small enough to justify the local nature of these potentials [cf. (56)].

VI. Conclusions

We considered the problem of ensuring communication integrity in networks of mobile robots. Unlike existing approaches that associate network and graph connectivity, network integrity is defined as the ability of a network to support desired communication rates. Our approach relies on introducing weights on the communication links to capture channel reliabilities, which then allows us to model routing by means of optimization problems that accept link reliabilities as inputs. The key idea proposed in this work was to join control of mobility and communications in a hybrid scheme with the discrete-time routing variables being the switching signal in the continuous-time motion controllers. We provided communication guarantees within a bounded error of optimality, which we explicitly characterized in terms of problem specific constants. This work focused on local convergence guarantees in the joint communication and mobility spaces and points to a new direction in systems and control theory on the interface with wireless networking.

Appendix
Proof of Theorem 1

The proof of Theorem 1 hinges on the use of the triangle inequality on the triangle with vertices $\boldsymbol{\lambda}(t_{n+1})$, $\boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})}$ and $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$ to bound the distance $\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})}\|$ as

$$\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})}\| \leq \|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\| + \|\boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})} - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\| \quad (58)$$

The bound in (58) separates consideration into the term $\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|$ and the term $\|\boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})} - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|$. The term $\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|$ is the distance between the iterate $\boldsymbol{\lambda}(t_{n+1})$ at time t_{n+1} and the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$ at time t_n. This distance we expect to be smaller than the corresponding distance $\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|$ at time t_n, because the dual update in (26) implements gradient descent in the dual function. This is the factor that results in an exponential convergence towards a neighborhood of the optimal. The term $\|\boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})} - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|$ captures the shift of the optimal multiplier as the team reconfigures. This is the factor that results in a steady state performance penalty. The proof of Theorem 1 relies in the bounding of the two terms in the right hand side of (58) that we respectively derive in lemmas 1 and 2.

Lemma 1: Consider team configurations $\mathbf{x}$ and $\mathbf{y}$ with corresponding reliabilities $R(\mathbf{x}_i, \mathbf{x}_j)$ and $R(\mathbf{y}_i, \mathbf{y}_j)$ and denote as $\boldsymbol{\lambda}^*_{\mathbf{x}}$ and $\boldsymbol{\lambda}^*_{\mathbf{y}}$ the optimal dual variables associated with the respective rate optimization problems in (25). Assume (A1) and (A4) are true and that reliabilities $R(\mathbf{x}_i, \mathbf{x}_j)$ and $R(\mathbf{y}_i, \mathbf{y}_j)$ do not differ by more than a given constant $\delta > 0$, i.e.,

$$|R(\mathbf{x}_i, \mathbf{x}_j) - R(\mathbf{y}_i, \mathbf{y}_j)| \leq \delta. \quad (59)$$

Then, the 2-norm distance between optimal multipliers satisfies

$$\left\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\right\|^2 \leq \frac{2\lambda_{\max} J}{m}\delta. \quad (60)$$

Proof: As per hypothesis, the dual function $g_{\mathbf{x}}(\boldsymbol{\lambda})$ associated with robot positions $\mathbf{x}$ is strongly convex. Hence, write the strong convexity inequality in (27) for $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*_{\mathbf{x}}$ and $\boldsymbol{\mu} = \boldsymbol{\lambda}^*_{\mathbf{y}}$

$$g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) \geq g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) + \nabla g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})^T(\boldsymbol{\lambda}^*_{\mathbf{y}} - \boldsymbol{\lambda}^*_{\mathbf{x}}) + \frac{m}{2}\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|^2. \quad (61)$$

Since $\boldsymbol{\lambda}^*_{\mathbf{x}}$ is the optimal argument of the convex function $g_{\mathbf{x}}(\boldsymbol{\lambda})$ we know that for all feasible $\boldsymbol{\lambda} \geq \mathbf{0}$ it must be $\nabla g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})^T(\boldsymbol{\lambda}^*_{\mathbf{y}} - \boldsymbol{\lambda}^*_{\mathbf{x}}) \geq 0$ from where it follows that,

$$g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) \geq g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) + \frac{m}{2}\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|^2. \quad (62)$$

The function $g_{\mathbf{y}}(\boldsymbol{\lambda})$ associated with configuration $\mathbf{y}$ is also strongly convex. An argument analogous to the one leading to (62) can be used to yield the bound

$$g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) \geq g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) + \frac{m}{2}\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|^2. \quad (63)$$

Combine the inequalities in (62) and (63) and reorder terms to obtain the bound

$$\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|^2 \leq \frac{1}{m}\left[\Big(g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) - g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})\Big) + \Big(g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) - g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{y}})\Big)\right]. \quad (64)$$

The bounds in (62)–(64) relate the distance $\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|$ to the differences in the dual values $g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) - g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})$ and $g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) - g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{y}})$ corresponding to the same functions evaluated at different dual arguments. Rearranging (64) yields

$$\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|^2 \leq \frac{1}{m}\left[\Big(g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) - g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{y}})\Big) + \Big(g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) - g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})\Big)\right]. \quad (65)$$

We interpret (65) as a bound on $\|\boldsymbol{\lambda}^*_{\mathbf{x}} - \boldsymbol{\lambda}^*_{\mathbf{y}}\|$ based on the differences in dual values $g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{y}}) - g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{y}})$ and $g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) - g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}})$, corresponding to *different functions* evaluated at the *same dual arguments*. These latter differences can be bounded in terms of problem constants.

To do so consider the optimal multiplier $\boldsymbol{\lambda}^*_{\mathbf{x}}$ of the dual function $g_{\mathbf{x}}(\boldsymbol{\lambda})$ associated with robot positions $\mathbf{x}$. Define $\mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}})$ and $\mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}})$ be the primal Lagrangian maximizers associated with robot positions $\mathbf{y}$ and multipliers $\boldsymbol{\lambda}^*_{\mathbf{x}}$. We can then write the value $g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}})$ of the dual function corresponding to spatial configuration $\mathbf{y}$ and multipliers $\boldsymbol{\lambda}^*_{\mathbf{x}}$ as

$$g_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) = f_0\Big(\mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}), \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}})\Big) + {\boldsymbol{\lambda}^*_{\mathbf{x}}}^T\Big[\mathbf{A}_{\mathbf{y}}\mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}}) - \mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}^*_{\mathbf{x}})\Big]. \quad (66)$$

For the same multiplier $\boldsymbol{\lambda}_{\mathbf{x}}^*$ consider the value of the dual function $g_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)$ associated with positions $\mathbf{x}$. This value satisfies

$$\begin{aligned} g_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{x}}^*) &= \max_{\mathbf{r},\mathbf{t}} \Big\{ f_0(\mathbf{r},\mathbf{t}) + \boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{x}}\mathbf{t} - \mathbf{r}] \Big\} \\ &\geq f_0\Big(\mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*), \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)\Big) + \boldsymbol{\lambda}_{\mathbf{x}}^{*T}\Big[\mathbf{A}_{\mathbf{x}}\mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*) - \mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)\Big], \end{aligned} \tag{67}$$

where the inequality follows because the Lagrangian maximization yields a value that is, at least, equal to the Lagrangian evaluated at $\mathbf{r}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)$, $\mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)$.

Subtracting (67) from (66) yields

$$g_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*) - g_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{x}}^*) \leq \boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\, \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*). \tag{68}$$

Likewise, consider the optimal multiplier $\boldsymbol{\lambda}_{\mathbf{y}}^*$ of the dual function $g_{\mathbf{y}}(\boldsymbol{\lambda})$ associated with robot positions $\mathbf{y}$ and let $\mathbf{r}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)$ and $\mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)$ be the primal Lagrangian maximizers associated with robot positions $\mathbf{x}$ and multipliers $\boldsymbol{\lambda}_{\mathbf{y}}^*$. Mimicking steps (66) through (68) we can write the analogous bound

$$g_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*) - g_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{y}}^*) \leq \boldsymbol{\lambda}_{\mathbf{y}}^{*T} [\mathbf{A}_{\mathbf{x}} - \mathbf{A}_{\mathbf{y}}]\, \mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*). \tag{69}$$

Substituting the bounds in (68) and (69) into the bound in (65) followed by application of the triangle inequality leads to

$$\begin{aligned} \|\boldsymbol{\lambda}_{\mathbf{x}}^* - \boldsymbol{\lambda}_{\mathbf{y}}^*\|^2 &\leq \frac{1}{m}\Big[\boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\, \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*) \\ &\qquad + \boldsymbol{\lambda}_{\mathbf{y}}^{*T} [\mathbf{A}_{\mathbf{x}} - \mathbf{A}_{\mathbf{y}}]\, \mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)\Big] \\ &\leq \frac{1}{m}\Big[\Big|\boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\, \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)\Big| \\ &\qquad + \Big|\boldsymbol{\lambda}_{\mathbf{y}}^{*T} [\mathbf{A}_{\mathbf{x}} - \mathbf{A}_{\mathbf{y}}]\, \mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)\Big|\Big]. \end{aligned} \tag{70}$$

To complete the proof we need to bound $\boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\, \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)$ and $\boldsymbol{\lambda}_{\mathbf{y}}^{*T} [\mathbf{A}_{\mathbf{x}} - \mathbf{A}_{\mathbf{y}}]\, \mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)$ in terms J, δ, and $\lambda_{\max}$. This requires some manipulation but is conceptually straightforward. We first observe that for any $\mathbf{t}$ it must be $|[\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t}| \leq (J\delta)\mathbf{1}$, where $\mathbf{1}$ is the column vector of all entries equal to one and the absolute value is taken on each element of the vector $[\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t}$. Indeed, the i-th component of the product $\mathbf{A}_{\mathbf{y}}\mathbf{t}$ is, by definition, the right hand side of the constraint in (3) evaluated at $\mathbf{t}$:

$$\begin{aligned} ([\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t})_i &= \sum (R(\mathbf{x}_i, \mathbf{x}_j) - R(\mathbf{y}_i, \mathbf{y}_j)) T_{ij} \\ &\quad - \sum (R(\mathbf{x}_j, \mathbf{x}_i) - R(\mathbf{y}_j, \mathbf{y}_i)) T_{ji} \end{aligned} \tag{71}$$

Since each of the differences $R(\mathbf{x}_i, \mathbf{x}_j) - R(\mathbf{y}_i, \mathbf{y}_j)$ are absolutely bounded by δ it follows that

$$|[\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t}|_j \leq \delta \sum T_{ij} + \delta \sum T_{ji} \leq J\delta. \tag{72}$$

To obtain the second inequality in (72) observe that the first sum contains all the T_{ij} transmission probabilities outgoing from robot i and it therefore satisfies $\sum T_{ij} \leq 1$. The second sum involves the $J-1$ probabilities of transmission to i. Since each of them is smaller than 1, we have $\sum T_{ji} \leq J-1$.

The bound in (72) can be written in vector form as $|[\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t}| \leq (J\delta)\mathbf{1}$. Thus, for any $\boldsymbol{\lambda}$ it must be

$$\big|\boldsymbol{\lambda}^T [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\mathbf{t}\big| \leq (J\delta)\boldsymbol{\lambda}^T\mathbf{1} \leq J\delta\lambda_{\max}, \tag{73}$$

were in the second equality we used the fact that because all components of $\boldsymbol{\lambda}$ are nonnegative, $\boldsymbol{\lambda}^T\mathbf{1} = \|\boldsymbol{\lambda}\|_1 \leq \lambda_{\max}$ as per Assumption (A4). The bound in (73) can be substituted for the terms $\big|\boldsymbol{\lambda}_{\mathbf{x}}^{*T} [\mathbf{A}_{\mathbf{y}} - \mathbf{A}_{\mathbf{x}}]\, \mathbf{t}_{\mathbf{y}}(\boldsymbol{\lambda}_{\mathbf{x}}^*)\big|$ and $\big|\boldsymbol{\lambda}_{\mathbf{y}}^{*T} [\mathbf{A}_{\mathbf{x}} - \mathbf{A}_{\mathbf{y}}]\, \mathbf{t}_{\mathbf{x}}(\boldsymbol{\lambda}_{\mathbf{y}}^*)\big|$ of (70) to obtain (60). ∎

Lemma 2: Let $\mathbf{x}(t_n)$ denote the team configuration at time t_n and $\boldsymbol{\lambda}_{\mathbf{x}(t_n)}^*$ the corresponding optimal dual variable. Consider the multiplier $\boldsymbol{\lambda}(t_n)$ at time t and the update $\boldsymbol{\lambda}(t_{n+1})$ obtained from $\boldsymbol{\lambda}(t_n)$ through (13). Assume the step size is bounded as $\epsilon \leq 1/M$ and define $\beta^2 \triangleq 1/(1+m\epsilon)$. Then, the distances between dual iterates and the optimal multiplier satisfy

$$\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}_{\mathbf{x}(t_n)}^*\|^2 \leq \beta^2 \|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}_{\mathbf{x}(t_n)}^*\|^2 \tag{74}$$

Proof: The derivation of (74) follows along the lines of the convergence proof of projected gradient descent algorithms; see e.g., [39, Ch. 3]. Define $S(\boldsymbol{\lambda}(t_n)) \triangleq -(1/\epsilon)(\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}(t_n))$ so that we can write the dual update as $\boldsymbol{\lambda}(t_{n+1}) = \boldsymbol{\lambda}(t_n) - \epsilon S(\boldsymbol{\lambda}(t_n))$. Commence the proof by recalling that as a consequence of the mean value theorem if the convex function $g_{\mathbf{x}}(\boldsymbol{\lambda})$ has Lipschitz gradients as stated in (28) it must be

$$g_{\mathbf{x}}(\boldsymbol{\mu}) - g_{\mathbf{x}}(\boldsymbol{\lambda}) \leq \nabla g_{\mathbf{x}}(\boldsymbol{\lambda})^T(\boldsymbol{\mu} - \boldsymbol{\lambda}) + \frac{M}{2}\|\boldsymbol{\mu} - \boldsymbol{\lambda}\|^2 \tag{75}$$

Use this expression for $\mathbf{x} = \mathbf{x}(t_n)$, $\boldsymbol{\mu} = \boldsymbol{\lambda}(t_{n+1})$, and $\boldsymbol{\lambda} = \boldsymbol{\lambda}(t_n)$ and reorder terms to bound $g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1}))$ as

$$\begin{aligned} g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) &\leq g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n)) \\ &\quad -\epsilon \nabla g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n))^T S(\boldsymbol{\lambda}(t_n)) + \frac{M\epsilon^2}{2}\|S(\boldsymbol{\lambda}(t_n))\|^2 \end{aligned} \tag{76}$$

where we used the definition of $S(\boldsymbol{\lambda}(t_n))$ to write $\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}(t_n) = -\epsilon S(\boldsymbol{\lambda}(t_n))$.

We first work on bounding the term $\nabla g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n))^T S(\boldsymbol{\lambda}(t_n))$. For doing so, consider variable $\boldsymbol{\lambda}$ whose components are not necessarily positive and recall that $\mathbb{P}[\boldsymbol{\lambda}]$ denotes the Euclidean projection of $\boldsymbol{\lambda}$ on the convex set $\boldsymbol{\lambda} \geq \mathbf{0}$. An important property of Euclidean projections in a convex set is that for any $\boldsymbol{\lambda}_0 \geq \mathbf{0}$ it holds

$$(\boldsymbol{\lambda} - \mathbb{P}[\boldsymbol{\lambda}])^T(\boldsymbol{\lambda}_0 - \mathbb{P}[\boldsymbol{\lambda}]) \leq 0. \tag{77}$$

Use this property with $\boldsymbol{\lambda} = \boldsymbol{\lambda}(t_n) - \epsilon\nabla g(\boldsymbol{\lambda}(t_n))$ and $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}_{\mathbf{x}(t_n)}^*$. Notice that in this case $\mathbb{P}[\boldsymbol{\lambda}] = \mathbb{P}[\boldsymbol{\lambda}(t_n) - \epsilon\nabla g(\boldsymbol{\lambda}(t_n))] = \boldsymbol{\lambda}(t_{n+1})$ by (26), which we can also write as $\mathbb{P}[\boldsymbol{\lambda}] = \boldsymbol{\lambda}(t_n) - \epsilon S(\boldsymbol{\lambda}(t_n))$ using the definition of $S(\boldsymbol{\lambda}(t_n))$. Thus, $\boldsymbol{\lambda} - \mathbb{P}[\boldsymbol{\lambda}] = \epsilon S(\boldsymbol{\lambda}(t_n)) - \epsilon\nabla g(\boldsymbol{\lambda}(t_n))$ from where it follows that (77) implies

$$\Big(S(\boldsymbol{\lambda}(t_n)) - \nabla g(\boldsymbol{\lambda}(t_n))\Big)^T \Big(\boldsymbol{\lambda}_{\mathbf{x}(t_n)}^* - \boldsymbol{\lambda}(t_n) + \epsilon S(\boldsymbol{\lambda}(t_n))\Big) \leq 0 \tag{78}$$

Expanding the inner product and rearranging terms in (78) yields

$$-\epsilon S(\boldsymbol{\lambda}(t_n))^T \nabla g(\boldsymbol{\lambda}(t_n)) \leq \nabla g(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) - S(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) - \epsilon\|S(\boldsymbol{\lambda}(t_n))\|^2 \quad (79)$$

Substituting the inequality in (79) into the one in (76) we obtain

$$\begin{aligned} g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) &\leq g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n)) \\ &+ \nabla g(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) \\ &- S(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) \\ &+ \left(\frac{M\epsilon^2}{2} - \epsilon\right)\|S(\boldsymbol{\lambda}(t_n))\|^2 \end{aligned} \quad (80)$$

Further using the fact that

$$g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_n)) + \nabla g(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) \leq g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)})$$

we can write

$$\begin{aligned} g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) &\leq g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) - \\ &- S(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)} - \boldsymbol{\lambda}(t_n)) \\ &+ \left(\frac{M\epsilon^2}{2} - \epsilon\right)\|S(\boldsymbol{\lambda}(t_n))\|^2. \end{aligned} \quad (81)$$

As required by hypothesis $\epsilon < 1/M$, which allows us to bound $M\epsilon^2/2 - \epsilon \leq -\epsilon/2$. Substituting this bound in (81) yields

$$\begin{aligned} &g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) \\ &\quad\leq g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) \\ &\qquad + S(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) - \frac{\epsilon}{2}\|S(\boldsymbol{\lambda}(t_n))\|^2. \end{aligned} \quad (82)$$

Consider now the squared distance to the optimum $\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2$ at iteration t_{n+1}. Recall that we defined $S(\boldsymbol{\lambda}(t_n))$ so as to write $\boldsymbol{\lambda}(t_{n+1}) = \boldsymbol{\lambda}(t_n) - \epsilon S(\boldsymbol{\lambda}(t_n))$ and expand the square to obtain

$$\begin{aligned} &\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2 \\ &\quad= \|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2 \\ &\qquad - 2\epsilon S(\boldsymbol{\lambda}(t_n))^T(\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) \\ &\qquad + \epsilon^2\|S(\boldsymbol{\lambda}(t_n))\|^2 \end{aligned} \quad (83)$$

The last two summands of the right hand sides of (82) and (83) coincide except for a -2ϵ factor. Solving (83) for these summands and substituting the result in (82) yields

$$\begin{aligned} &g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) - g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) \\ &\quad\leq \frac{1}{2\epsilon}\left(\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2 - \|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2\right). \end{aligned} \quad (84)$$

Finally, note that because $g_{\mathbf{x}}(\boldsymbol{\lambda})$ is a convex function it holds $\nabla g_{\mathbf{x}}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)})^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}) \geq 0$ for any $\boldsymbol{\lambda}$. Using this property in the strong convexity assumption in (27) for $\boldsymbol{\mu} = \boldsymbol{\lambda}(t_{n+1})$ and $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}$ we obtain

$$\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\|^2 \leq \frac{2}{m}\left(g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}(t_{n+1})) - g_{\mathbf{x}(t_n)}(\boldsymbol{\lambda}^*_{\mathbf{x}(t_n)})\right). \quad (85)$$

Substituting (84) into (85), regrouping terms and using the definition of $\beta^2 = 1/(1 + m\epsilon)$ the result in (74) follows. ■

To finalize the proof of Theorem 1 just notice that the second term in the right hand side of (58) satisfies the bound in Lemma 1, while the square of the first term satisfies the bound in Lemma 2. Therefore

$$\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})}\| \leq \beta\|\boldsymbol{\lambda}(t_n) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_n)}\| + \sqrt{\frac{2\lambda_{\max}J}{m}}\delta \quad (86)$$

Applying the inequality in (86) recursively between times t_{n+1} and t_0 yields

$$\|\boldsymbol{\lambda}(t_{n+1}) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_{n+1})}\| \leq \beta^{n+1}\|\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}^*_{\mathbf{x}(t_0)}\| + \sum_{u=0}^{n}\beta^u\sqrt{\frac{2\lambda_{\max}J}{m}}\delta \quad (87)$$

Observing that $\sum_{u=0}^{n}\beta^u \leq \sum_{u=0}^{\infty}\beta^u = 1/(1-\beta)$ the result in (32) follows after shifting time indices from t_{n+1} to t_n.

REFERENCES

[1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[2] W. Ren and R. W. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[3] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[4] E. M. Royer and C.-K. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Personal Commun.*, vol. 6, no. 2, pp. 46–55, Apr. 1999.

[5] I. Stojmenovic, A. Nayak, and J. Kuruvila, "Design guidelines for routing protocols in ad hoc and sensor networks with a realistic physical layer," *IEEE Commun. Mag.*, vol. 43, pp. 101–106, Mar. 2005.

[6] J. Kuruvila, A. Nayak, and I. Stojmenovic, "Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 6, pp. 1267–1275, Jun. 2005.

[7] A. Neskovic, N. Neskovic, and G. Paunovic, "Modern approaches in modeling of mobile radio systems propagation environment," *IEEE Commun. Surveys*, vol. 3, no. 3, pp. 1–12, 2000.

[8] J. D. Parsons, *The Mobile Radio Propagation Channel*. Hoboken, NJ: Wiley, 2000.

[9] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*. Hoboken, NJ: Wiley, 1995.

[10] H. Ando, Y. Oasa, I. Suzuki, and M. Yamashita, "Distributed memoryless point convergence algorithm for mobile robots with limited visibility," *IEEE Trans. Robot. Automat.*, vol. 15, no. 5, pp. 818–828, Oct. 1999.

[11] Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph laplacian," *IEEE Trans. Autom. Control*, vol. 51, no. 1, pp. 116–120, Jan. 2006.

[12] M. C. DeGennaro and A. Jadbabaie, "Decentralized control of connectivity for multi-agent systems," in *Proc. 45th IEEE Conf. Decision Control*, San Diego, CA, Dec. 2006, pp. 3628–3633.

[13] M. Ji and M. Egerstedt, "Coordination control of multi-agent systems while preserving connectedness," *IEEE Trans. Robot.*, vol. 23, no. 4, pp. 693–703, Aug. 2007.

[14] M. M. Zavlanos and G. J. Pappas, "Potential fields for maintaining connectivity of mobile networks," *IEEE Trans. Robot.*, vol. 23, no. 4, pp. 812–816, Aug. 2007.

[15] M. Schuresko and J. Cortes, "Distributed motion constraints for algebraic connectivity of robotic networks," *J. Intelligent Robotic Syst.*, vol. 56, no. 1–2, pp. 99–126, Sep. 2009.
[16] M. M. Zavlanos and G. J. Pappas, "Distributed connectivity control of mobile networks," *IEEE Trans. Robot.*, vol. 24, no. 6, pp. 1416–1428, Dec. 2008.
[17] E. Stump, A. Jadbabaie, and V. Kumar, "Connectivity management in mobile robot teams," in *Proc. IEEE Int. Conf. Robot. Automat.*, Pasadena, CA, May 2008, pp. 1525–1530.
[18] H. Lundgren, E. Nordstrom, and C. Tschudin, "The gray zone problem in ieee 802.11b based ad hoc networks," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 6, no. 3, pp. 104–105, Jul. 2002.
[19] D. DeCouto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multihop wireless routing," in *Proc. of Int. ACM Conf. Mobile Comput. Networking*, San Diego, CA, Sep. 2006, pp. 134–146.
[20] A. Ribeiro, Z.-Q. Luo, N. D. Sidiropoulos, and G. B. Giannakis, "Modelling and optimization of stochastic routing for wireless multihop networks," in *Proc. 26th Ann. Joint Conf. IEEE Comput. Commun. Societies (INFOCOM)*, Anchorage, AL, May 2007, pp. 1748–1756.
[21] A. Ribeiro, N. D. 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.
[22] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Operational Res. Society*, vol. 49, no. 3, pp. 237–252, 1998.
[23] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, vol. 22, pp. 28–43, Feb. 2002.
[24] R. Srikant, *The Mathematics of Internet Congestion Control*, 1st ed. Boston, MA: Birkhauser, 2004.
[25] A. Eryilmaz and R. Srikant, "Joint congestion control, routing, and mac for stability and fairness in wireless networks," *IEEE J. Selected Areas in Commun.*, vol. 24, no. 8, pp. 1514–1524, Aug. 2006.
[26] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Selected Areas in Commun.*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.
[27] Y. Yi and S. Shakkottai, "Hop-by-hop congestion control over a wireless multi-hop network," *IEEE/ACM Trans. Network.*, vol. 15, no. 1, pp. 1548–1559, Feb. 2007.
[28] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Selected Areas in Commun.*, vol. 23, no. 1, pp. 89–103, Jan. 2005.
[29] M. J. Neely, "Energy optimal control for time-varying wireless networks," *IEEE Trans. Infor. Theory*, vol. 52, no. 7, pp. 2915–2934, Jul. 2006.
[30] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136–1144, Jul. 2004.
[31] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6369–6386, Dec. 2010.
[32] A. Ribeiro and G. B. Giannakis, "Separation principles in wireless networking," *IEEE Trans. Infor. Theory*, vol. 56, no. 9, pp. 4488–4505, Sep. 2010.
[33] A. Ribeiro and G. B. Giannakis, "Optimal fdma over wireless fading mobile ad-hoc networks," in *Proc. IEEE ICASSP*, Las Vegas, NV, Mar. 2008, pp. 2765–2768.
[34] M. J. Neely, "Universal scheduling for networks with arbitrary traffic, channels, and mobility," in *Proc. 49th IEEE Conf. Decision Control*, Altanta, GA, Dec. 2010, pp. 1822–1829.
[35] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions*. Berlin, Germany: Springer-Verlag, 1985.
[36] E. Rimon and D. Koditschek, "Robot navigation functions on manifolds with boundary," *Adv. Appl. Mathemat.*, vol. 11, pp. 412–442, 1990.
[37] E. Rimon and D. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Trans. Robot. Automat,*, vol. 8, no. 5, pp. 501–518, Oct. 1992.
[38] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level measurements from an 802.11b mesh network," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, pp. 121–132, Oct. 2004.
[39] D. Bertsekas and J. N. Tsitsiklis, *Nonlinear Programming*, 2nd ed. Cambridge, MA: Athena Scientific, 1999.

Michael M. Zavlanos (S'05–M'09) received the Diploma in mechanical engineering from the National Technical University of Athens, Athens, Greece, in 2002, and the M.S.E. and Ph.D. degrees in electrical and systems engineering from the University of Pennsylvania, Philadelphia, in 2005 and 2008, respectively.

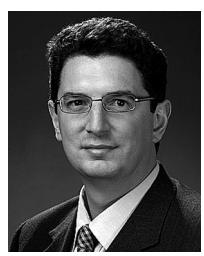
From 2008 to 2009 he was a Post-Doctoral Researcher in the Department of Electrical and Systems Engineering at the University of Pennsylvania, Philadelphia. He then joined the Stevens Institute of Technology, Hoboken, NJ, as an Assistant Professor of Mechanical Engineering, where he remained until 2012. He is currently an Assistant Professor of Mechanical Engineering and Materials Science at Duke University, Durham, NC. His current research interests include a wide range of topics in the emerging discipline of networked systems and science, with applications in robotic, sensor, biomolecular, and social networks. He is particularly interested in hybrid solution techniques, on the interface of control theory with the discrete science of networks and graphs.

Dr. Zavlanos was a recipient of the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award in 2011 and was a finalist for the Best Student Paper Award at the 45th IEEE Conference on Decision and Control in 2006.

Alejandro Ribeiro (S'02–M'07) received the B.Sc. in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, in 1998, and the M.Sc. and Ph.D. degrees in electrical engineering in 2005 and 2007, respectively, from the University of Minnesota, Minneapolis.

From 1998 to 2003 he was a Member of the Technical Staff at Bellsouth Montevideo. From 2003 to 2008, he was with the Department of Electrical and Computer Engineering, the University of Minnesota, Minneapolis. He is currently an Assistant Professor at the Department of Electrical and Systems Engineering at the University of Pennsylvania, Philadelphia. His research interests lie in the areas of communication, signal processing, and networking. His current research focuses on the study of networked phenomena arising in technological, human, and natural networks.

Dr. Ribeiro received the 2012 S. Reid Warren, Jr. Award presented by Penn's undergraduate student body for outstanding teaching, the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award in 2010, and student paper awards at ICASSP 2005 and ICASSP 2006. He is also a Fulbright Scholar.



George J. Pappas (S'90–M'91–SM'04–F'09) received the Ph.D. degree in electrical engineering and computer sciences from the University of California, Berkeley, in 1998.

He is currently the Joseph Moore Professor of Electrical and Systems Engineering at the University of Pennsylvania, Philadelphia. He is a member of the General Robotics, Automation, Sensing and Perception (GRASP) Laboratory and serves as the Deputy Dean for Research in the School of Engineering and Applied Science. His current research interests include hybrid and embedded control systems, hierarchical control systems, distributed control systems, nonlinear control systems, with applications to robotics, unmanned aerial vehicles, biomolecular networks, and green buildings.

Dr. Pappas has received numerous awards, including the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award in 2002, the NSF Presidential Early Career Award for Scientists and Engineers in 2002, the 2009 George S. Axelby Outstanding Paper Award, the Eliahu Jury Award for Excellence in Systems Research from the University of California, Berkeley, and the 2010 Antonio Ruberti Outstanding Young Researcher Prize.