For presentation in *CDC 2015*

On Distributed Optimization using Generalized Gossip

Zhanhong Jiang† Soumik Sarkar† Kushal Mukherjee‡
zhjiang@iastate.edu soumiks@iastate.edu MukherK@utrc.utc.com
† Department of Mechanical Engineering, Iowa State University, Ames, IA 50011, USA
‡ United Technologies Research Center, Cork, Ireland

***Abstract*— This paper presents a generalized gossip-based algorithm to solve distributed optimization problems in multi-agent networks, especially for multiple supply-demand optimization problems. The proposed algorithm provides a generalization such that the optimization process can operate in the entire spectrum of "complete consensus" to "complete disagreement". A user-defined control parameter θ is identified for controlling such tradeoff as well as the temporal convergence properties. Analytical results for first moment convergence analysis are presented and it is shown that with $\theta \to 0$, the formulation boils down to a classical consensus based protocol. Beyond the control parameter, the agent interaction matrix is also shown to be useful for effectively suppressing large localized uncertainties in subgradient estimation. A practical use case regarding building zone temperature control is presented as a numerical example for validation.**

1. Introduction

In recent times, large-scale networks have received considerable interest from industry and academia due to their impact in areas such as robotics, intelligent surveillance and reconnaissance, transportation networks and smart buildings [1], [2], [3]. One challenge that is faced by these large-scale sensor and actuator networks is to autonomously optimize the behaviors of agents or allocate resources within these networks while reducing the computational and economic cost. State-of-the-art techniques employ cooperative and non-cooperative [4] distributed optimization [5] to obtain the best agent behavior and/or resource allocation. In these methods, a global decision (scalar/vector quantity) is shared by all agents to prevent local minima [6]. Subgradient methods [7], [8], [9], [10], [11], [12] are widely used to iteratively refine the estimates of the shared decision for each agent in a distributed manner. In [13], a scheme combining consensus algorithm with subgradient method was described for solving the convex optimization problems. In [3], a theoretical framework was established for consensus and cooperation in networked multi-agent systems.

Generalized gossip algorithms [14], [15], [16] essentially extend gossip algorithms [17], [18] to cases where perfect consensus is not required to be archived. A tradeoff between the decision propagation radius (i.e., how far a decision spreads from its source) and localization of information is chosen to vary the dissemination of agent beliefs throughout the network. While consensus of the belief state of agents is not guaranteed, proximal agents are more likely to share similar beliefs. Therefore, in sensor networks, generalized gossip aims at describing the observed phenomenon at more 'local' level as compared to pure consensus algorithms. The objective of this paper is to combine the generalized gossip algorithm with subgradient approach and present the generalized gossip-based subgradient algorithm developed for distributed optimization. The optimal solution is derived from performing a more 'local' consensus of behavior or allocation of resources. With a user-defined generalizing parameter in the agent interaction policy, the trade-off between propagation radius and localization gradient may be controlled to yield a spectrum of optimal solutions ranging from globally optimal solution (complete consensus) to greedy locally optimal solution (no compromise).

The specific contributions of this paper beyond the existing work are:(1) formulation of a new technique for distributed optimization by incorporating the generalized gossip algorithm into a subgradient optimization framework, (2) first moment convergence analysis of consensus and/or discord between agents for distributed optimization, (3) proposition of agent interaction matrix adaptation for suppressing localized large uncertainties in subgradient estimation, (4) validation of the proposed algorithm in a simulation test bed for temperature control in building zones.

2. Background and Problem Setup

Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ consisting of N agents, where $\mathcal{V} = \{1, 2, ..., N\}$ and $\mathcal{A} \subseteq \mathcal{V} \times \mathcal{V}$. If $(i, j) \in \mathcal{A}$, then agent i can communicate with agent j. Let a distributed optimization problem be defined on the network as follows:

$$\begin{aligned} &\text{minimize } , f(x) = \sum_{i=1}^{N} f^i(x) \\ &\text{subject to } , x \in \mathbb{X} \end{aligned} \tag{1}$$

where $f^i : \mathbb{R}^M \longrightarrow \mathbb{R}$ are agent level objective functions (possibly convex or non-convex), $\mathbb{X}$ is a nonempty, closed, and compact subset of $\mathbb{R}^M$. x is a vector whose i^{th} component is represented by x^i. While in this initial paper, only unconstrained case is considered, constrained case will be an important future work.

A. Preliminary Background

The basic definitions [7], [13] and assumptions used in this paper are:

Definition 1: A vector $g \in \mathbb{R}^M$ is a subgradient of a convex function $f : \mathbb{R}^M \longrightarrow \mathbb{R}$ at a point $z \in \mathbb{R}^M$ if

$$f(y) \geq f(z) + g^T(y - z), \forall y \in \mathbb{R}^M \tag{2}$$

Definition 2: The set of all subgradients of a convex function of f at $z \in \mathbb{R}^M$ is called the subdifferential of f at z, and is denoted by $\partial f(z)$:

$$\partial f(z) = \{g \in \mathbb{R}^M | f(y) \geq f(z) + g^T(y - z), \forall y \in \mathbb{R}^M\} \tag{3}$$

Definition 3: $\exists \epsilon \geq 0$, for all $y \in \mathbb{R}^M$, if $f(y) \geq f(z) + g^T(y-z)-\epsilon$, then g is a ϵ-subgradient. Namely, $g \in \partial_\epsilon f(z)$.

Note, scalar optimization variables can be easily converted to M dimensional vectors as suggested in the definitions presented above. For notational simplifications scalar agent level optimization variables x^i are used without loss of generality.

Assumption 1 (Subgradient boundedness): There exists a scalar G for all $i = 1, ..., N$ such that

$$\|g^i(x)\|_2 \leq G, \forall g^i(x) \in \partial f^i(x), \forall x \in \mathbb{X} \tag{4}$$

Assumption 2: The optimal solution set x^* is nonempty.

The standard distributed subgradient formulation for multi-agent distributed optimization (see [8]) is as follows:

$$x^i(k + 1) = \sum_{j=1}^{m} a^i_j(k)x^j(k) - \alpha^i(k)d^i(k) \tag{5}$$

where a^i_j are weights, $\alpha^i(k)$ are step sizes, $d^i(k)$ are the subgradients of f^i at $x^i(k)$ and k is a time index. Depending on choice weights, this formulation has been shown to solve multi-agent optimization problems primarily via consensus. Choice of step size is also key to obtain desired convergence properties.

B. Generalized Gossip Algorithm

"Agreement" or "consensus" is one of the most widely used concepts in the area of multi-agent systems for information fusion, decision-making, information propagation and distributed optimization. Among various protocols, gossip-based consensus algorithms are quite popular due to their simple nature yet powerful properties as well as strong analytical results [19], [20]. Recently, in the context of distributed information propagation in a mobile sensor network [21], the present authors proposed a generalized gossip protocol [14]. One of the key observations made in the study was that a user defined parameter is able to control the fundamental tradeoff between information propagation radius and localization gradient. This paper applies this formulation to a distributed optimization problem where the tradeoff is between "global optimal" achieved through compromise and "local optimal" (s) achieved in a greedy manner by individual agents.

The generalized gossip protocol outlined in [14] for belief propagation in a mobile sensor network is as follows:

$$\upsilon^i_\theta(k + 1) = (1 - \theta) \sum_{j\in\{i\}\cup Nb(i)} \pi_{ij}(k)\upsilon^j_\theta(k) + \theta\chi^i(k) \tag{6}$$

where $\theta \in (0, 1]$ is a user-defined control parameter, $Nb(i)$ is the set of agents in the neighborhood of agent i. In this set, agents can communicate with agent i during the time span between k and $k+1$. π_{ij} is the element of i-th row and j-th column in the agent interaction matrix $\Pi \in \mathbb{R}^N \times \mathbb{R}^N$. υ and χ are the vectors representing agent belief measure and state characteristics function (quantifying observations made by agents), respectively. While the agent interaction matrix may in general be time varying, it maintains the *doubly stochastic* property.

3. Generalized Gossip-based Subgradient Algorithm

In this section, the proposed generalized gossip-based subgradient algorithm is proposed for distributed optimization. In this formulation, the discrete-time update law (derived from eqn. 5 and 6) for the optimization variable is as follows:

$$x^i(k+1) = (1-\theta) \sum_{j\in\{i\}\cup Nb(i)} \pi_{ij}(k)x^j(k)+\theta(x^i(k)-\nabla^i(k)) \tag{7}$$

where $\nabla^i(k)$ is the subgradient of f^i at $x^i(k)$ computed by agent i. Note, there is no explicit stepsize parameter in this update rule. However, user defined parameter θ acts as the weight on the subgradient term at every update step.

In a vector notation, the update rule becomes:

$$x(k + 1) = (1 - \theta)\Pi(k)x(k) + \theta(x(k) - \nabla(k)). \tag{8}$$

Note, as the control parameter θ approaches 0, this policy boils down to a standard consensus-based optimization protocol using subgradients. On the other hand, as θ approaches 1, interaction among agents reduce significantly and individual agent-level optimization variable tend to converge to their respective local minima based on individual subgradients. Formal discussion on this aspect will follow the analytical results presented in the sequel.

This initial paper presents analytical results for the first moment along with generic numerical results. Second moment analysis remains a critical future work. Following are notations and lemmas required for the convergence analysis.

For first moment analysis, ensemble average (over agents) of $x(k)$ and $\nabla(k)$ are denoted by $\bar{x}(k)$ and $\bar{\nabla}(k)$ respectively. They are defined as:

$$\bar{x}(k) = \frac{1}{N}\mathbf{1}x(k) = \sum_{i=1}^{N} x^i(k) \tag{9}$$

$$\bar{\nabla}(k) = \frac{1}{N}\mathbf{1}\nabla(k) = \sum_{i=1}^{N}\nabla^i(k) \quad (10)$$

where $\mathbf{1}$ is a row vector with all elements being 1.

Note, multiplying the update rule described in eqn. 8 by $\frac{1}{N}\mathbf{1}$ yields the following relationship (as Π is doubly stochastic):

$$\bar{x}(k+1) = \bar{x}(k) - \theta\bar{\nabla}(k) \quad (11)$$

The form of this equation is similar to that of the classical subgradient method.

Next, optimal function values are denoted by f^*, that are assumed to be finite. Without loss of generality the optimal set is represented by x^*, i.e., $x^* = \{x \in \mathbb{R}| \sum_{i=1}^{N} f^i(x) = f^*\}$. Note, optimal points can be multi-dimensional and not necessarily have to be scalar. Here are a few necessary lemmas and proposition required for convergence analysis.

Lemma 1: If $\|x^i(k) - x^j(k)\|_2 \leq \sigma < \infty, \forall i, j = 1, ..., N$, then $\|x^i(k) - \bar{x}(k)\|_2 \leq \sigma$. See [13] for proof.

Note, this lemma suggests that if pairwise distance between optimization variables of any two agents is bounded then distance between the average and any point is bounded by the same quantity. Before presenting the next lemma, the following proposition is required.

Proposition 1: If *Assumption 1* holds, then for a sequence $\{\bar{\nabla}(k)\}$, $\forall k$,

$$\|\bar{\nabla}(k)\|_2 \leq G.$$

Proof: $\|\bar{\nabla}(k)\|_2 = \|\frac{1}{N}\sum_{i=1}^{N}\nabla^i(k)\|_2 \leq \frac{1}{N}[\|\nabla^1(k)\|_2 + \|\nabla^2(k)\|_2 + \cdots + \|\nabla^N(k)\|_2] \leq \frac{1}{N}(NG) = G$. ■

The proposition suggests that if all individual agent level subgradient norms are bounded, then norm of the average subgradient is also bounded by the same quantity.

Lemma 2: If *Assumption 1* holds, then, for a sequence $\{\bar{x}(k)\}$, $\forall k$ and $z \in \mathbb{R}$,

$$f(z) \geq f(\bar{x}(k)) + N\bar{\nabla}(k)(z - \bar{x}(k)) - \epsilon \quad (12)$$

where $\epsilon = 2NG\sigma$.

Proof: Using *Definition 1* and *Assumption 1*

$$\begin{aligned} f^i(x^i(k)) &\geq f^i(\bar{x}(k) + \bar{\nabla}^i(k)(x^i(k) - \bar{x}(k)) \\ &\geq f^i - \|\bar{\nabla}^i(k)\|_2\|x^i(k) - \bar{x}(k)\|_2 \\ &\geq f^i(\bar{x}(k)) - G\sigma. \end{aligned} \quad (13)$$

Using the above relationship for all N agents

$$f(x^i(k)) \geq f(\bar{x}(k)) - NG\sigma. \quad (14)$$

For any $z \in \mathbb{R}$, there exists

$$\begin{aligned} f^i(z) &\geq f^i(x^i(k)) + \nabla^i(k)(z - x^i(k)) \\ &\geq f^i(x^i(k)) + \nabla^i(k)(z - \bar{x}(k) + \bar{x}(k) - x^i(k)) \\ &\geq f^i(x^i(k)) + \nabla^i(k)(z - \bar{x}(k) - \nabla^i(k)) \\ &- \|\nabla^i(k)\|_2\|\bar{x}(k) - x^i(k)\|_2 \\ &\geq f^i(x^i(k)) + \nabla^i(k)(z - \bar{x}(k) - \nabla^i(k)) - G\sigma. \end{aligned} \quad (15)$$

Again, using the above relationship for all N agents

$$f(z) \geq f(x^i(k)) + \sum_{i=1}^{N}\nabla^i(k)(z - \bar{x}(k)) - NG\sigma. \quad (16)$$

Substituting eqn. 14 into eqn. 16, the following inequality can be obtained

$$f(z) \geq f(\bar{x}(k)) + \sum_{i=1}^{N}\nabla^i(k)(z - \bar{x}(k)) - 2NG\sigma. \quad (17)$$

Combining eqn. 17 and eqn. 10, the desired result is achieved. ■

Note, *Lemma 2* shows that $N\bar{\nabla}$ is an ϵ-subgradient of $\bar{x}$.

4. Statistical Convergence Analysis

In this section, first moment convergence analysis for the generalized gossip based subgradient algorithm is presented. The analysis begins with the following lemma.

Lemma 3: If *Assumptions 1, 2* holds, then, for a sequence $\{\bar{x}(k)\}$, $\forall k$,

$$\begin{aligned} \|\bar{x}(k+1) - x^*\|_2^2 \leq \|\bar{x}(k) - x^*\|_2^2 - 2\theta\frac{1}{N}(f(\bar{x}(k)) \\ - f^*) + 4\theta G\sigma + \theta^2 G^2. \end{aligned} \quad (18)$$

Proof: By using Lemma 2,

$$\begin{aligned} \|\bar{x}(k+1) - x^*\|_2^2 &= \|\bar{x}(k) - \theta\bar{\nabla}(k) - x^*\|_2^2 \\ &= \|\bar{x}(k) - x^*\|_2^2 - 2\theta\bar{\nabla}(k) \\ &(\bar{x}(k) - x^*) + \theta^2\|\bar{\nabla}(k)\|_2^2 \\ &\leq \|\bar{x}(k) - x^*\|_2^2 \\ &- 2\theta\frac{1}{N}(f(\bar{x}(k)) - f^*) \\ &+ 4\theta G\sigma + \theta^2 G^2. \end{aligned} \quad (19)$$

where $f^* = f(x^*)$. The last inequality follows from $f(x^*) \geq f(\bar{x}(k)) + N\bar{\nabla}(k)(x^* - \bar{x}(k)) - 2NG\sigma$ and the *Assumption 2*. ■

Lemma 3 suggests that $\bar{x}$ gets closer to x^* with every iteration when $f(\bar{x})$ is much greater than f^*. However, as $f(\bar{x})$ comes very close to f^*, $\bar{x}$ is not guaranteed to get closer to x^*. This is due to the two positive terms in the equation, $4\theta G\sigma$ and $\theta^2 G^2$. If the control parameter θ is reduced, then the positive terms reduce in magnitude and $\bar{x}$ approaches x^*.

Next, *Theorem 1* is introduced to provide an upper bound for the optimized function value.

Theorem 1: If *Assumptions 1, 2* holds, then, for a sequence $\{x^i(k)\}$, $\forall k$ and $i = 1, \ldots, N$,

$$\begin{aligned} f^* \leq f(x^i(k))_{min} \leq f^* + \frac{N\|\bar{x}(1) - x^*\|_2^2}{2m\theta} \\ +3NG\sigma + \frac{N\theta G^2}{2} \end{aligned} \quad (20)$$

where $f(x^i(k))_{min} = \min\{f(x^i(1)), \ldots, f(x^i(m))\}$, m is the number of iterations, G is the upper bound of subgradients, σ is the upper bound of Euclidean distance between

$x^i(k)$ and $\bar{x}(k)$ and θ is the control parameter. As the subgradient is not always in the descent direction, the best value need to be tracked that best approaches f^*. Hence, after m iterations, the best value is denoted by $f(x^i(k))_{min}$.

Proof: The lower bound is easily obtained when the optimal solution set exists. Therefore, the proof mainly deals with the upper bound. Recalling *Lemma 3*,

$$\begin{aligned}\|\bar{x}(k+1)-x^*\|_2^2 &\leq \|\bar{x}(k)-x^*\|_2^2\\ &- 2\theta\frac{1}{N}(f(\bar{x}(k))-f^*)+4\theta G\sigma+\theta^2G^2,\end{aligned} \tag{21}$$

then, applying this inequality recursively for m iterations

$$\begin{aligned}\|\bar{x}(k+1)-x^*\|_2^2 &\leq \|\bar{x}(1)-x^*\|_2^2\\ &- 2\theta\frac{1}{N}\sum_{k=1}^{m}(f(\bar{x}(k))-f^*)+4\theta mG\sigma\\ &+m\theta^2G^2.\end{aligned} \tag{22}$$

Since $\|\bar{x}(k+1)-x^*\|_2^2 \geq 0$, the inequality above can be written as

$$\begin{aligned}0 &\leq \|\bar{x}(1)-x^*\|_2^2\\ &- 2\theta\frac{1}{N}\sum_{k=1}^{m}(f(\bar{x}(k))-f^*)+4\theta mG\sigma\\ &+m\theta^2G^2.\end{aligned} \tag{23}$$

Let $f(\bar{x}(k))_{min} = \min\{f(\bar{x}(1)),\ldots,f(\bar{x}(m))\}$. Then, the second term on the right hand side of the above equation satisfies

$$\begin{aligned}&2\theta\frac{1}{N}\sum_{k=1}^{m}(f(\bar{x}(k))-f^*)\\ &\geq (\sum_{k=1}^{m}2\theta\frac{1}{N})\min_{k=1,\ldots,m}(f(\bar{x}(k))-f^*).\end{aligned} \tag{24}$$

Combining the above two equations another inequality follows

$$\begin{aligned}f(\bar{x}(k))_{min}-f^* &= \min_{k=1,\ldots,m} f(\bar{x}(k))-f^*\\ &\leq \frac{N\|\bar{x}(1)-x^*\|_2^2}{2m\theta}+2NG\sigma+\frac{N\theta G^2}{2}.\end{aligned} \tag{25}$$

The goal is to obtain the inequality in terms of the sequence $\{x^i(k)\}$. Hence, the relationship between $f(x^i(k))$ and $f(\bar{x}(k))$ would be required.

By recalling that

$$f^i(\bar{x}(k)) \geq f^i(x^i(k))+\nabla^i(\bar{x}(k)-x^i(k)),$$

the following expression can be obtained using *Assumption 1*,

$$f(x^i(k)) \leq f(\bar{x}(k))+NG\sigma. \tag{26}$$

Finally, substituting eqn. 26 into eqn. 25, the upper bound described in Theorem is obtained. ■

Theorem 1 demonstrates the convergence of function values and it can be concluded that for a given θ, as number of iteration m increases, the upper bound converges to $f^*+3NG\sigma+\frac{N\theta G^2}{2}$. Thus, $f(x^i(k))_{min}$ converges within $3NG\sigma+\frac{N\theta G^2}{2}$ from the optimal value. Also, note that effect of initial condition dies out with a large value of m.

Next, the convergence characteristics are analyzed as the control parameter approaches extreme values 0 or 1.

Let a sequence $\{\theta_k\}$, $k=1,2,\ldots,m$ follow the *diminishing step size* properties.

There exists an integer N_1, that satisfies $\theta_k \leq \frac{\delta}{NG^2}, \delta>0, \forall k>N_1$. Then there exists another integer N_2 such that

$$\sum_{k=1}^{m}\theta_k \geq \frac{1}{\delta}(N\|\bar{x}(1)-x^*\|_2^2+NG^2\sum_{k=1}^{N_1}\theta_k^2), \forall m>N_2.$$

This inequality holds because $\lim_{m\to\infty}\sum_{k=1}^{m}\theta_k=\infty$.

Now, let $\mathfrak{N}=max\{N_1,N_2\}$, and the following *Theorem 2* can be stated.

Theorem 2: If *Assumptions 1 and 2* holds, then, for a sequence $\{x^i(k)\}$, if θ_k satisfies *diminishing step size*, $\forall k$, and $i=1,\ldots,N$,

$$f(x^i(k))_{min} \leq f^*+\delta+3NG\sigma, \forall m>\mathfrak{N}. \tag{27}$$

Proof: By replacing θ with θ_k in eqn. 20 and rewriting it,

$$\begin{aligned}&f(x^i(k))_{min}-f^* \leq\\ &\frac{N\|\bar{x}(1)-x^*\|_2^2+6N\sum_{k=1}^{m}\theta_kG\sigma+NG^2\sum_{k=1}^{m}\theta_k^2}{2\sum_{k=1}^{m}\theta_k},\end{aligned} \tag{28}$$

then rewriting the above inequality,

$$\begin{aligned}&f(x^i(k))_{min}-f^* \leq\\ &\frac{N\|\bar{x}(1)-x^*\|_2^2+NG^2\sum_{k=1}^{N_1}\theta_k^2}{2\sum_{k=1}^{m}\theta_k}\\ &+\frac{6N\sum_{k=1}^{m}\theta_kG\sigma}{2\sum_{k=1}^{m}\theta_k}\\ &+\frac{NG^2\sum_{k=N_1+1}^{m}\theta_k^2}{2\sum_{k=1}^{N_1}\theta_k+2\sum_{k=N_1+1}^{m}\theta_k}\\ &\leq \frac{N\|\bar{x}(1)-x^*\|_2^2+NG^2\sum_{k=1}^{N_1}\theta_k^2}{\frac{2}{\delta}[N\|\bar{x}(1)-x^*\|_2^2+NG^2\sum_{k=1}^{N_1}\theta_k^2]}\\ &+3NG\sigma+\frac{NG^2\sum_{k=N_1+1}^{m}\frac{\delta}{NG^2}\theta_k}{2\sum_{k=N_1+1}^{m}\theta_k}\\ &=\delta+3NG\sigma.\end{aligned} \tag{29}$$

Since δ is arbitrarily small, this implies that if the iteration number $m\to\infty$, $\lim_{k\to\infty}f(x^i(k)) \leq f^*+3NG\sigma$. ■

This result shows that as $\theta\to 0$, the proposed formulation boils down to the standard subgradient formulation as analyzed in [13]. With small values of G and σ, the gap between the upper bound and the true optimal value reduces.

Now, the case where $\theta\to 1$ is analyzed by presenting the following theorem.

Theorem 3: If *Assumptions1, 2* holds, then, for a sequence $\{x^i(k)\}$, with $\theta \to 1$, $\forall k$ and $i = 1, \ldots, N$,

$$f(x^i(k))_{min} \leq f^* + \frac{N\|\bar{x}(1) - x^*\|_2^2}{2m} + 3NG\sigma + \frac{NG^2}{2}. \tag{30}$$

Proof: Directly follows from *Theorem 1*.

Now, with iteration number $m \to \infty$:

$$\lim_{k\to\infty} f(x^i(k)) \leq f^* + 3NG\sigma + \frac{NG^2}{2}. \blacksquare \tag{31}$$

Note, that the gap between the upper bound and the true optimal value increases as θ increases from 0 and approaches 1. Furthermore, with $\theta = 1$, the update rule boils down to: $x^i(k+1) = x^i(k) - \nabla^i(k) \ \ \forall i$. It follows from this equation that with $\theta = 1$, individual agents reach their own locally optimal values as there is no interaction among agents.

Remark 4.1 *First moment analysis presented in this paper shows that while lower θ reduces the error with respect to the global optimal value, larger θ increases that. This is further explained by the observation that with $\theta = 1$, the distributed optimization problem turns into separate optimization problems for individual agents. Hence, θ controls the tradeoff between "degree of consensus" and "degree of disagreement".*

Remark 4.2 *Although it is important to consider subgradients of all agents (with low value of θ), highly uncertain subgradient computation by certain agent(s) can reduce the effectiveness of the optimization algorithm. However, in case Π is controllable such that appropriate Π can be chosen online while maintaining certain constraints such as its stochastic nature, one can potentially adapt Π to suppress the impact of large uncertainty stemming from a particular agent. In this context, the role of agent interaction matrix is numerically explored in the following section for suppressing large uncertainties in subgradient computations.*

5. NUMERICAL EXAMPLE

This section presents numerical results to validate the proposed generalized gossip distributed optimization algorithm in the context of optimizing supply air temperature for minimizing energy consumption in a building involving ten zones (agents) while achieving their respective comfort requirements. In this problem, a general heating, ventilation, and air-conditioner (HVAC) system associated with a building is investigated. In such a system, zone temperature regulation is implemented typically through the supply air flow provided by a central air handling unit (AHU). Please see [22] for further details. By using the proposed algorithm, the global supply air temperature gets updated following the information exchange among zones.

Figure 1 shows that supply air temperature for each zone converge to the optimized value of $29.9°$C as a consequence of a small value of the control parameter choice ($\theta = 0.05$). In this example, the network is fully connected, that is, every zone exchanges information with every other to cooperatively come up with the desired supply air temperature. As θ increases from 0, "degree of disagreement" increases as suggested in the previous section and shown in fig. 2. Temporal behavior of convergence (i.e., convergence gets slower with lower value of θ) can also be observed. With mid ranges of θ, an interesting "clustering" phenomenon can be observed where zones with similar requirements converge to similar supply air temperature requirements. This suggests that the proposed generalized gossip-based distributed optimization policy can be very useful in the context of multi resource and consumption entities or multiple supply-demand optimization problems.

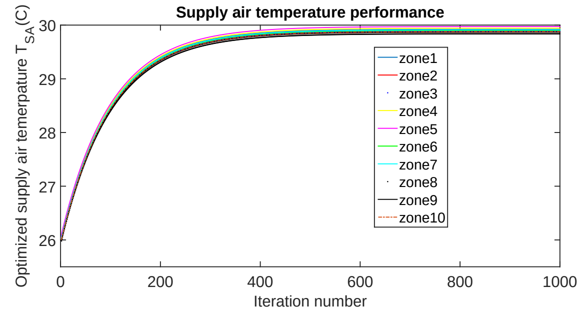

Fig. 1. Supply air temperature convergence plots with iteration number using $\theta = 0.05$; Optimal value $29.9°$C

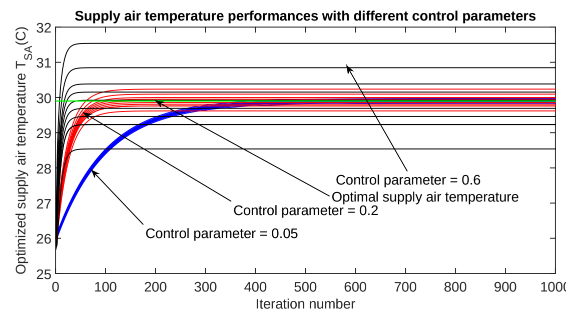

Fig. 2. Supply air temperature optimization performance with different values of control parameter θ; ; Optimal value $29.9°$C; while lower θ increases *degree of consensus* convergence time increases as well

Figure 3 shows the optimization performance with uniform agent interaction matrix Π (fully connected network) and it is evident that the optimal supply air temperature performance suffers due to the large uncertainty in test zone 10. On the other hand, use of a nonuniform Π (non-fully connected network) that almost isolates zone 10 reduces its impact on the overall optimization problem (see fig. 4). However, its own performance suffers as a consequence. In both cases, the control parameter used was $\theta = 0.1$. Therefore, there are two types of control actions that can be taken in the proposed framework, namely the control parameter θ and

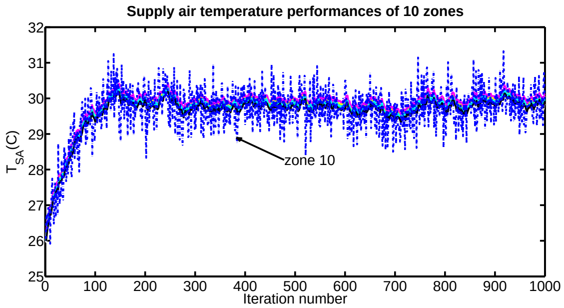

Fig. 3. Supply air temperature optimization performance with uniform agent interaction matrix under uncertainties in subgradient computation

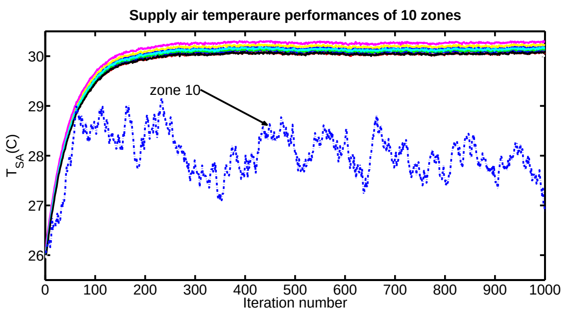

Fig. 4. Supply air temperature optimization performance with nonuniform agent interaction matrix under uncertainties in subgradient computation

the agent interaction matrix Π.

6. CONCLUSIONS AND FUTURE WORKS

In this paper, a generalized gossip-based subgradient algorithm is proposed to solve distributed optimization problems in multi-agent networks. Analytical results for first moment convergence analysis are presented and it is shown that with $\theta \to 0$, the formulation boils down to a classical consensus based protocol. While θ controls this tradeoff, it also controls the temporal convergence properties. A practical use case regarding building zone temperature control is presented as a numerical example to illustrate the proposed algorithm. Beyond the control parameter θ, the agent interaction matrix Π can also be used to effectively suppress large uncertainties in subgradient estimation stemming from certain agents.

While the ongoing effort is focusing on second moment convergence analysis, some other potential future research directions are: Effectively using "clustering" effects (in mid ranges of θ) for set point optimization at multiple resources; quantifying and propagation analysis of uncertainties stemming from subgradient computations; extending to constrained optimization problems; validation on real-life large scale supply-demand networks.

REFERENCES

[1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor," *Automatical Control, IEEE Transactions on*, vol. 48, 2006.

[2] S. Patterson, B. Bamieh, and A. El Abbadi, "Convergence rates of distributed average consensus with stochastic link failures," *Automatic Control, IEEE Transactions on*, vol. 55, pp. 880–892, April 2010.

[3] R. O. Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, January, 2007.

[4] E. Altman, K. Avrachenkov, G. Miller, and B. Prabhu, "Discrete power control: Cooperative and non-cooperative optimization," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pp. 37–45, May 2007.

[5] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *Automatical Control, IEEE Transactions on*, vol. 31, pp. 803–812, September, 1986.

[6] V. Borkar and P. Varaya, "Asymptotic agreement in distributed estimation," *Automatical Control, IEEE Transactions on*, vol. 27, pp. 650–655, June, 1982.

[7] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," 2003.

[8] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *Automatic Control, IEEE Transactions on*, vol. 54, pp. 48–61, January, 2009.

[9] A. Nedic and D. Bertsekas, "Convergence rate of incremental subgradient algorithms," *Stochastic Optimization: Algorithms and Applications*, pp. 263–304, 2001.

[10] D. Yuan, S. Xu, and H. Zhao, "Distributed primal-dual subgradient method for multiagent optimization via consensus algorithm," *Systems, Man, and Cybernetics, IEEE Transactions on*, vol. 41, pp. 1715–1724, December, 2011.

[11] Q. Long, C. Wu, and X. Wang, "A system of nonsmooth equations solver based upon subgradient method," *Applied Mathematics and Computation*, pp. 284–299, 2015.

[12] A. Nedic and A. Ozdaglar, "Subgradient methods in network resource allocation: rate analysis," *in Proceedins of 42nd Annual Conference on Information Sciences and Systems, Princeton, NJ*, March, 2008.

[13] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," *in Proceedings of the 47th IEEE Conference on Decision and Control, Caucun, Mexico*, pp. 4185–4190, December,2008.

[14] S. Sarkar, K. Mukherjee, and A. Ray, "Distributed decision propagation in mobile-agent proximity networks," *International Journal of Control*, vol. 86, no. 6, pp. 1118–1130, 2013.

[15] S. Sarkar and K. Mukherjee, "Event-triggered decision propagation in proximity networks," *Frontiers in Robotics and AI*, vol. 1, p. 15, 2014.

[16] P. Chattopadhyay, D. K. Jha, S. Sarkar, and A. Ray, "Path planning in GPS-denied environments: A collective intelligence approach," in *Proceedings of American Control Conference, Chicago, IL, USA*, July 2015.

[17] S. S. Ram, A. Nedic, and V. V. Veeravalli, "Asynchronous gossip algorithms for stochastic optimization," *in Proceedings of Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, China*, December, 2009.

[18] J. Lu, C. Y. Tang, P. R. Regier, and T. D. Bow, "A gossip algorithm for convex consensus optimization over networks," *in Proceedings of American Control Conference, Marriott Waterfront, Baltimore, MD*, pp. 301–308, June - July, 2010.

[19] D. Shah, "Gossip algorithms," *Foundations and Trends in Networking*, vol. 3, pp. 1–125, 2008.

[20] A. Salehi Tahbaz and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *Automatical Control, IEEE Transactions on*, vol. 55, pp. 225–230, 2010.

[21] S. Sarkar, K. Mukherjee, A. Srivastav, and A. Ray, "Distributed decision propagation in mobile agent networks," in *Proceedings of Conference on Decision and Control, Atlanta, GA*, 2010.

[22] Y. Ma, G. Anderson, and F. Borrelli, "A distributed predictive control approach to building temperature regulation," *in Proceedings of 2011 American Control Conference, San Francisco, California*, June-July,2011.