Distributed Solution for a Maximum Variance Unfolding Problem with Sensor and Robotic Network Applications

Andrea Simonetto, Tamás Keviczky, and Dimos V. Dimarogonas

Abstract—We focus on a particular non-convex networked optimization problem, known as the Maximum Variance Unfolding problem and its dual, the Fastest Mixing Markov Process problem. These problems are of relevance for sensor networks and robotic applications. We propose to solve both these problems with the same distributed primal-dual subgradient iterations whose convergence is proven even in the case of approximation errors in the calculation of the subgradients. Furthermore, we illustrate the use of the algorithm for sensor network applications, such as localization problems, and for mobile robotic networks applications, such as dispersion problems.

I. INTRODUCTION

We are interested in solving a particular non-convex problem, known as the Maximum Variance Unfolding (MVU) problem with the intention to devise globally optimal distributed algorithms. The MVU problem can be described as the problem of choosing a set of points to be as far apart as possible from each other, measured by their variance, while respecting local maximum distance constraints. This problem arises in different research areas, such as localization [1], mechanics [2], linear algebra [3], and unsupervised (machine) learning [4]. Together with its dual (known as the Fastest Mixing Markov Process problem, or FMMP problem), it is linked to important research questions in networked optimization, such as the maximization and minimization of the algebraic connectivity of the underlying graph under given constraints.

In this paper we solve both the MVU problem and its dual, the FMMP problem, with a primal-dual subgradient algorithm using a fixed step-size. Convergence of the primal-dual scheme is proven even in the case when errors are present in the computation of the subgradients. This situation happens often in practice when the subgradient determination involves iterative schemes. The proposed primal-dual algorithm is then applied to relevant sensor and mobile network applications, among which localization and dispersion problems.

This paper is organized as follows. Section II describes the MVU problem, Section III presents its dual, the FMMP problem, and the basic ingredients for the proposed distributed algorithm of Section IV. Section V extends the obtained results for the case in which errors are present in the determination of the subgradients. In Section VI a numerical example is presented to analyze the performance of the proposed algorithms.

A. Simonetto and T. Keviczky are with the Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands {a.simonetto, t.keviczky}@tudelft.nl
D. V. Dimarogonas is with the Automatic Control group, School of Electrical Engineering, Royal Institute of Technology (KTH), 100 44 Stockholm, Sweden, dimos@kth.se

Sections VII and VIII describe two applications, an anchor-free localization problem in sensor networks and a dispersion problem in mobile robotic networks. Section IX concludes the paper and presents some future research directions.

II. PROBLEM FORMULATION

We use the standard notation $\mathbf{1}_n$ and I_n to indicate a vector of all ones of dimension n and an identity matrix of dimensions $n \times n$, respectively. The notation $\langle \cdot, \cdot \rangle$ is the dot product while $||\cdot||$ is the 2-norm. For a real symmetric matrix A, the notation $A \succ 0$ indicates that A is positive definite, $\lambda_i(A)$ denotes its i-th eigenvalue whose eigenvector is $\mathbf{v}_i$.

We study constrained optimization problems on a network of computing nodes. The network is modeled as a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with vertices (nodes) in the set $\mathcal{V} = \{1, \ldots, N\}$ and pairs of nodes as edges in the set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. We denote the cardinality of $\mathcal{E}$ as E, the set of neighbors of node i as $\mathcal{N}_i = \{j | (i, j) \in \mathcal{E}\}$, while L is the Laplacian of the graph $\mathcal{G}$, [5]. Let $r_{ij} \in \mathbb{R}$ be a "distance" bound associated with the edge (i, j) of the graph $\mathcal{G}$ connecting the nodes.

We are interested in solving the following non-convex problem

$$\underset{x_1, \ldots, x_N}{\text{maximize}} \quad \sum_{i=1}^{N} ||x_i||^2 \tag{1a}$$

$$\text{subject to} \quad ||x_i - x_j||^2 \leq r_{ij}^2 \quad \forall (i, j) \in \mathcal{E} \tag{1b}$$

$$\sum_{i=1}^{N} x_i = 0 \tag{1c}$$

where each variable $x_i \in \mathbb{R}$ is associated to node i, while we denote with $x \in \mathbb{R}^N$ the vector collecting all the optimization variables $x = (x_1, \ldots, x_N)^\top$. (In Section VII we will consider an extension to $x_i \in \mathbb{R}^n$).

Problem (1) is known as the Maximum Variance Unfolding problem, or MVU [2], [4]. Under the assumption that the communication graph $\mathcal{G}$ is connected, the MVU problem (1) has a (possibly not unique) optimal solution x^{opt}, which makes all the inequality constraints active [2].

Although the MVU problem (1) is non-convex, it is well-known that it can be relaxed into a convex problem by the change of variables $X = xx^\top$. Pursuing this transformation

we arrive at the convex SDP[1] [2], [3]

$$\underset{X}{\text{maximize}} \qquad \text{trace}\,(X) \qquad\qquad \text{(2a)}$$

$$\text{subject to} \qquad X_{ii} + X_{jj} - X_{ij} - X_{ji} \le r_{ij}^2, \qquad \text{(2b)}$$
$$\forall (i,j) \in \mathcal{E}$$
$$\mathbf{1}_N^\top X \mathbf{1}_N = 0, \quad X \succeq 0 \qquad\qquad \text{(2c)}$$

The non-convex problem (1) and the convex formulation (2) are equivalent in the sense that they yield the same optimal value, i.e., $x^{\text{opt}\top} x^{\text{opt}} = \text{trace}\,(X^{\text{opt}})$. For this reason, in many applications it is often convenient to transform the non-convex (1) into the convex (2), which can be solved efficiently [1], [4].

Both the non-convex problem (1) and the convex one (2) can be solved with centralized algorithms (although, for the non-convex case the algorithms may lead to local optima). Furthermore, methods have been proposed to approximate the convex problem (2) in order to reduce its computational complexity when the size of the problem becomes large (e.g., $N > 1000$), see [1].

III. DISTRIBUTED SOLUTIONS

Solving the two problems (1) and (2) in a distributed way is more challenging. In principle, the non-convex (1) could be solved using a Sequential Quadratic Programming approach [6], [7], and the resulting quadratic programs distributed among the nodes with some decomposition methods. In practice, at best, this type of approach leads only to a local optimum. On the other hand, the convex formulation (2) is constrained via matrix (in)equalities that are difficult to decouple among the nodes. In particular, the projection over the constraint $X \succeq 0$ would require the knowledge of the spectral decomposition of the full matrix X, see [8]. Although a distributed algorithm is available [9] to obtain such decomposition, it is limited to matrices that exhibit the same sparsity as the underlying graph and therefore it is not directly applicable in our case. Furthermore, the knowledge of X^{opt} would not automatically imply that x^{opt} can be computed in a distributed way.

It turns out however that the dualization of the convex problem (2) overcomes most of these issues and allows us to devise globally optimal distributed algorithms. This leads us to study the dual formulation of the MVU problem, which is discussed next.

Let $w_{ij} \in \mathbb{R}_+$ be a weight associated with the edge ij, while let $w \in \mathbb{R}_+^E$ be the stacked vector of all the weights. Let $L(w)$ be the Laplacian matrix constructed with these weights w_{ij} and let $\lambda_2(w)$ be its second smallest eigenvalue (i.e., the algebraic connectivity). The dual of the SDP (2) can be written

---

[1]In the case of scalar x_i the convex SDP (2) is a rank 1 relaxation. This means that if X^{opt} happens to be of rank 1 then it is possible to reconstruct x^{opt}. In practice as explained in [4], although not formally guaranteed, the trace cost function naturally leads to low rank solutions for X.

as [2]

$$\underset{w}{\text{maximize}} \qquad \lambda_2(w) \qquad\qquad \text{(3a)}$$

$$\text{subject to} \qquad \sum_{(i,j)\in\mathcal{E}} r_{ij}^2 w_{ij} \le 1 \qquad \text{(3b)}$$
$$w \ge 0 \qquad\qquad \text{(3c)}$$

or in the equivalent form (since both the objective $\lambda_2(w)$ and the constraint function $\sum_{(i,j)\in\mathcal{E}} r_{ij}^2 w_{ij}$ are positive homogeneous)

$$\underset{w}{\text{minimize}} \qquad \sum_{(i,j)\in\mathcal{E}} r_{ij}^2 w_{ij} \qquad \text{(4a)}$$

$$\text{subject to} \qquad \lambda_2(w) \ge 1 \qquad\qquad \text{(4b)}$$
$$w \ge 0 \qquad\qquad\qquad \text{(4c)}$$

where the inequality $w \ge 0$ has to be interpreted element-wise. The problems (3) and (4) are convex, since $\lambda(w)$ is a concave function of the weights w [10]. Moreover, we remark that at optimality $\lambda_2(w^{\text{opt}}) = 1$.

Problem (3) is known as the Fastest Mixing Markov Process (FMMP) on a graph, which is the problem of determining the weights w that maximize the algebraic connectivity of the graph, under a certain linear bound on w. The FMMP problem can thus also be written as

$$\underset{w}{\text{minimize}} \qquad \sum_{(i,j)\in\mathcal{E}} r_{ij}^2 w_{ij} \qquad \text{(5a)}$$

$$\text{subject to} \qquad L(w) + (1/N)\mathbf{1}_N\mathbf{1}_N^\top \succeq I_N \qquad \text{(5b)}$$
$$w \ge 0 \qquad\qquad\qquad\qquad \text{(5c)}$$

where the Laplacian L depends linearly on the weights w.

Remark 1: Notice that we can substitute the constraint $\lambda_2(w) \ge 1$ with $\lambda_2(w) \ge \bar{\lambda}_2$ without any difficulty. In fact, due to the linearity of the Laplacian L on w, the scaling $\tilde{w} = w/\bar{\lambda}_2$ would normalize the problem to $\bar{\lambda}_2 = 1$.

Both problems (4) and (5) are duals of (2). In particular, assuming the presence of a Slater vector for (4), we can prove that the duality gap is zero [2], meaning

$$\sum_{(i,j)\in\mathcal{E}} r_{ij}^2 w_{ij}^{\text{opt}} = \text{trace}\,(X^{\text{opt}})$$

Furthermore when solving the convex MVU problem (2), at optimality its dual variables will be optimal for the FMMP problem (5) and vice versa. We remark that this is also true for the non-convex MVU problem (1) if its global maximum is found. We formalize these primal-dual relationships by the use of KKT optimal conditions. In particular, the optimal couple $(w^{\text{opt}}, X^{\text{opt}})$ satisfies:

- Primal-Dual feasibility:

$$w^{\text{opt}} \le 0, \quad L(w^{\text{opt}}) \succeq I_N - (1/N)\mathbf{1}_N\mathbf{1}_N^\top$$

$$\mathbf{1}_N^\top X^{\text{opt}} \mathbf{1}_N = 0, \quad X^{\text{opt}} \succeq 0,$$

$$X_{ii}^{\text{opt}} + X_{jj}^{\text{opt}} - X_{ij}^{\text{opt}} - X_{ji}^{\text{opt}} \le r_{ij}^2, \ \forall (i,j) \in \mathcal{E}$$

- Complementary slackness on edges:

$$(X_{ii}^{\mathrm{opt}} + X_{jj}^{\mathrm{opt}} - X_{ij}^{\mathrm{opt}} - X_{ji}^{\mathrm{opt}} - r_{ij}^2)w_{ij}^{\mathrm{opt}} = 0, \ \forall (i,j) \in \mathcal{E}$$

- Matrix complementary slackness:

$$L(w^{\mathrm{opt}})X^{\mathrm{opt}} = X^{\mathrm{opt}}$$

The last condition means that the range of X^{opt} lies in the eigenspace of $L(w^{\mathrm{opt}})$ associated with $\lambda_2(w^{\mathrm{opt}})$ (which we recall to be one). This leads to the following result.

Indicate with $\mathbf{v}_2^{\mathrm{opt}}$ the normalized eigenvector associated with $\lambda_2(w^{\mathrm{opt}})$ and call c^{opt} the optimal cost for the FMMP problem, i.e., $c^{\mathrm{opt}} = \sum_{(i,j) \in \mathcal{E}} r_{ij}^2 w_{ij}^{\mathrm{opt}}$, then an optimal solution for X is

$$X^{\mathrm{opt}} = c^{\mathrm{opt}} \mathbf{v}_2^{\mathrm{opt}} \mathbf{v}_2^{\mathrm{opt}\top} \tag{6}$$

Furthermore, if $\lambda_2(w^{\mathrm{opt}})$ is isolated, this solution is also unique [2] (in other words if $\lambda_2(w^{\mathrm{opt}})$ is isolated then X^{opt} is rank 1 and therefore can be used to compute x^{opt}). In this case the relation (6) also yields to

$$x^{\mathrm{opt}} = \sqrt{c^{\mathrm{opt}}} \mathbf{v}_2^{\mathrm{opt}} \tag{7}$$

for a global optimizer of the non-convex (1).

Equation (7) will be an important ingredient in the design of a distributed globally optimal algorithm for the non-convex problem (1) as we illustrate next.

IV. Proposed Distributed Algorithm

We are interested in solving the MVU problem (1) in a distributed fashion. In order to do so, we propose to utilize its dual convex FMMP (4) problem and a primal-dual subgradient technique. (We remark that $\lambda_2(w)$ is a non-smooth function of w). The intention is to design a globally optimal distributed algorithm for (1) by solving in a distributed way the convex FMMP (4).

First of all, we recall some useful definitions and some necessary preliminary results. Let f be a generic convex function $f : \mathbb{R}^n \to \mathbb{R}$, while let g be a generic concave function $g : \mathbb{R}^n \to \mathbb{R}$. The vector $\mathbf{s} \in \mathbb{R}^n$ is a subgradient of the convex function f at a point $x \in \mathbb{R}^n$ if

$$f(y) \geq f(x) + \langle \mathbf{s}, y - x \rangle, \quad \text{for all } y \in \mathbb{R}^n$$

while $\mathbf{s}$ is a subgradient[2] of the concave function g at a point $x \in \mathbb{R}^n$ if

$$g(y) \leq g(x) + \langle \mathbf{s}, y - x \rangle, \quad \text{for all } y \in \mathbb{R}^n$$

Given a strictly positive scalar $\varepsilon > 0$, the vector $\mathbf{s}$ is an ε-subgradient of f at a point $x \in \mathbb{R}^n$ if

$$f(y) \geq f(x) + \langle \mathbf{s}, y - x \rangle - \varepsilon, \quad \text{for all } y \in \mathbb{R}^n$$

while $\mathbf{s}$ is an ε-subgradient of g at a point $x \in \mathbb{R}^n$ if

$$g(y) \leq g(x) + \langle \mathbf{s}, y - x \rangle + \varepsilon, \quad \text{for all } y \in \mathbb{R}^n$$

---
[2] We note that some authors refer to $\mathbf{s}$ as supergradient if it is the subgradient of a concave function.

The concept of ε-subgradient is useful when the computation of the subgradient of a given function is affected by approximation errors. This is the reason why the schemes that employ ε-subgradients instead of subgradients are often referred to as approximate subgradient methods [11].

Consider the convex FMMP problem (4). Let the vector $\mathbf{q_{ij}} \in \mathbb{R}^N$ be

$$(\mathbf{q_{ij}})_q = \begin{cases} 1 & \text{if } q = i \\ -1 & \text{if } q = j \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

We remark that the subscripts $\mathbf{i}$ and $\mathbf{j}$ are bold since the object $\mathbf{q_{ij}}$ is a vector that refers to the nodes i and j and not the element (i,j) of $\mathbf{q}$.

The Laplacian $L(w)$ of the underlying graph $\mathcal{G}$ on which problem (4) is based upon can be written as

$$L(w) = \sum_{(i,j) \in \mathcal{E}} \mathbf{q_{ij}} \mathbf{q_{ij}}^\top w_{ij}$$

The algebraic connectivity of $L(w)$ is a concave function of w since for every $\tilde{w} \in \mathbb{R}_+^E$ [10]

$$\lambda_2(\tilde{w}) \leq \lambda_2(w) + \text{trace}\left(\langle \mathbf{v}_2 \mathbf{v}_2^\top, L(\tilde{w}) - L(w) \rangle\right)$$

where $\mathbf{v}_2$ is computed on $L(w)$. Substituting the expression of $L(w)$ we obtain

$$\lambda_2(\tilde{w}) \leq$$

$$\lambda_2(w) + \text{trace}\left(\left\langle \mathbf{v}_2 \mathbf{v}_2^\top, \sum_{(i,j) \in \mathcal{E}} \mathbf{q_{ij}} \mathbf{q_{ij}}^\top (\tilde{w}_{ij} - w_{ij}) \right\rangle\right) =$$

$$\lambda_2(w) + \text{trace}\left(\sum_{(i,j) \in \mathcal{E}} \mathbf{v}_2 \mathbf{v}_2^\top \mathbf{q_{ij}} \mathbf{q_{ij}}^\top (\tilde{w}_{ij} - w_{ij})\right) =$$

$$\lambda_2(w) + \sum_{(i,j) \in \mathcal{E}} \text{trace}\left(\mathbf{v}_2 \mathbf{v}_2^\top \mathbf{q_{ij}} \mathbf{q_{ij}}^\top\right)(\tilde{w}_{ij} - w_{ij}) =$$

$$\lambda_2(w) + \sum_{(i,j) \in \mathcal{E}} (\mathbf{v}_{2i} - \mathbf{v}_{2j})^2 (\tilde{w}_{ij} - w_{ij})$$

where $\mathbf{v}_{2i}$ and $\mathbf{v}_{2j}$ are the i-th and j-th component of $\mathbf{v}_2$. Therefore the vector $\nabla_w \lambda_2(w) \in \mathbb{R}^E$ of components

$$(\nabla_w \lambda_2(w))_{ij} = (\mathbf{v}_{2i} - \mathbf{v}_{2j})^2 \tag{9}$$

is a subgradient for (the concave) $\lambda_2(w)$ at w.

Let $\mathcal{L}(w, \mu)$ be the Lagrangian function associated with the FMMP problem (4), i.e.,

$$\mathcal{L}(w, \mu) = \sum_{(i,j) \in \mathcal{E}} r_{ij}^2 w_{ij} + \mu(1 - \lambda_2(w)) \tag{10}$$

where $\mu \in \mathbb{R}_+$ is the dual variable of w. Let $\nabla_w \mathcal{L}(w, \mu)$ and $\nabla_\mu \mathcal{L}(w, \mu)$ be the subgradients of $\mathcal{L}(w, \mu)$ with respect to w and μ, respectively. These subgradients can be expressed component-wise as

$$\begin{aligned} (\nabla_w \mathcal{L}(w, \mu))_{ij} &= r_{ij}^2 - \mu \left(\nabla_w \lambda_2(w)\right)_{ij} = \\ &= r_{ij}^2 - \mu(\mathbf{v}_{2i} - \mathbf{v}_{2j})^2 \\ \nabla_\mu \mathcal{L}(w, \mu) &= 1 - \lambda_2(w) \end{aligned}$$

In order to solve in a distributed way the FMMP problem (4), we consider the primal-dual iterations

$$w^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+}\left[w^{(\tau)} - \alpha\nabla_w\mathcal{L}(w^{(\tau)},\mu^{(\tau)})\right] \quad (11)$$

$$\mu^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+}\left[\mu^{(\tau)} + \alpha\nabla_\mu\mathcal{L}(w^{(\tau)},\mu^{(\tau)})\right] \quad (12)$$

with a constant step-size $\alpha > 0$ and where $\mathcal{P}_{\mathbb{R}_+}$ is the projection on the positive orthant.

We assume the following standard assumption on the boundedness of the subgradients.

Assumption 1: The subgradients $\nabla_w\mathcal{L}(w,\mu)$ and $\nabla_\mu\mathcal{L}(w,\mu)$ used in the method defined by (11)-(12) are uniformly bounded, i.e., there is a constant $\Lambda > 0$ such that

$$\max\left\{\left\|\nabla_w\mathcal{L}(w^{(\tau)},\mu^{(\tau)})\right\|, \left\|\nabla_\mu\mathcal{L}(w^{(\tau)},\mu^{(\tau)})\right\|\right\} \leq \Lambda$$

for all iterations $\tau \geq 0$.

We define the running averages $\bar{w}^{(\tau)}$ and $\bar{\mu}^{(\tau)}$ generated by:

$$\bar{w}^{(\tau)} = \frac{1}{\tau}\sum_{j=0}^{\tau-1} w^{(j)}, \quad \bar{\mu}^{(\tau)} = \frac{1}{\tau}\sum_{j=0}^{\tau-1} \mu^{(j)}$$

We can cite the following theorem from [12] that guarantees convergence of the couple $(\bar{w}^{(\tau)}, \bar{\mu}^{(\tau)})$ to a saddle-point of the Lagrangian (10).

Theorem 1: [Proposition 1 of [12]] Under Assumption 1 the following relations for the iterates (11)-(12) hold true:

(a) For all $\tau \geq 1$,

$$-\frac{||\mu^{(0)} - \mu^{\mathrm{opt}}||^2}{2\alpha\tau} - \frac{\alpha\Lambda^2}{2} \leq$$
$$\frac{1}{\tau}\sum_{j=0}^{\tau-1}\mathcal{L}(w^{(j)},\mu^{(j)}) - \mathcal{L}(w^{\mathrm{opt}},\mu^{\mathrm{opt}}) \leq$$
$$\frac{||w^{(0)} - w^{\mathrm{opt}}||^2}{2\alpha\tau} + \frac{\alpha\Lambda^2}{2}$$

(b) The averages $\bar{w}^{(\tau)}$ and $\bar{\mu}^{(\tau)}$ satisfy the following relation for all $\tau \geq 1$:

$$-\frac{||\mu^{(0)} - \mu^{\mathrm{opt}}||^2 + ||w^{(0)} - \bar{w}^{(\tau)}||^2}{2\alpha\tau} - \alpha\Lambda^2 \leq$$
$$\mathcal{L}(\bar{w}^{(\tau)},\bar{\mu}^{(\tau)}) - \mathcal{L}(w^{\mathrm{opt}},\mu^{\mathrm{opt}}) \leq$$
$$\frac{||w^{(0)} - w^{\mathrm{opt}}||^2 + ||\mu^{(0)} - \bar{\mu}^{(\tau)}||^2}{2\alpha\tau} + \alpha\Lambda^2$$

The result in (a) gives bounds on the averaged function values $\frac{1}{\tau}\sum_{j=0}^{\tau-1}\mathcal{L}(w^{(j)},\mu^{(j)})$ in terms of distances of the initial iterates $w^{(0)}$ and $\mu^{(0)}$ from w^{opt} and μ^{opt} that determine a saddle point for $\mathcal{L}(w,\mu)$. The averaged function values converge to the saddle point $\mathcal{L}(w^{\mathrm{opt}},\mu^{\mathrm{opt}})$ within error bound $\alpha\Lambda^2/2$ with convergence rate of $1/\tau$. The result in (b) provides bounds on the function value $\mathcal{L}(\bar{w}^{(\tau)},\bar{\mu}^{(\tau)})$ of the averaged iterates $\bar{w}^{(\tau)}$ and $\bar{\mu}^{(\tau)}$ in terms of distances of the averaged iterates from the initial iterates and saddle point vectors. Under Assumption 1, this result shows that the function values of the averaged iterates $\mathcal{L}(\bar{w}^{(\tau)},\bar{\mu}^{(\tau)})$ converge to the saddle-point value $(w^{\mathrm{opt}},\mu^{\mathrm{opt}})$ within error bound $\alpha\Lambda^2$ with convergence

rate of $1/\tau$. The error is produced by the use of a constant step-size and can be tuned by choosing a smaller step-size value at the price of increasing the number of iterations τ. As a result, Theorem 1 gives explicit trade-offs between accuracy (in terms of α) and computational complexity (in terms of τ) in choosing the step-size value.

In the following theorem we characterize the value of the optimal dual variable μ^{opt}, which will enable us to derive a global optimal optimizer for the non-convex MVU (1).

Theorem 2: The optimal value of the dual variable μ^{opt} is unique and equal to the cost of the FMMP problem (4), i.e., $\mu^{\mathrm{opt}} = c^{\mathrm{opt}}$.

Proof: An optimal point for the FMMP problem (4) is a fixed point of the iterations (11)-(12), due to Theorem 1. In particular, an optimizer of (4), must satisfy $\forall(i,j) \in \mathcal{E}$

$$w_{ij}^{\mathrm{opt}} = \mathcal{P}_{\mathbb{R}_+}\left[w_{ij}^{\mathrm{opt}} - \alpha(r_{ij}^2 - \mu^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2)\right]$$

since $\alpha > 0$, this means that: either

$$w_{ij}^{\mathrm{opt}} = 0 \text{ and } r_{ij}^2 - \mu^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2 > 0$$

or

$$w_{ij}^{\mathrm{opt}} > 0 \text{ and } r_{ij}^2 - \mu^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2 = 0.$$

Due to the fact that w^{opt} is a nonzero vector, we can write these alternatives compactly as

$$\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}\left(r_{ij}^2 - \mu^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2\right) = 0,$$

thus

$$\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}r_{ij}^2 = \mu^{\mathrm{opt}}\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2,$$

and therefore

$$\begin{aligned}
c^{\mathrm{opt}} &= \mu^{\mathrm{opt}}\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}(\mathbf{v}_{2i}^{\mathrm{opt}} - \mathbf{v}_{2j}^{\mathrm{opt}})^2 \\
&= \mu^{\mathrm{opt}}\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}\mathrm{trace}\left(\mathbf{v}_2^{\mathrm{opt}}\mathbf{v}_2^{\mathrm{opt}\top}\mathbf{q_{ij}}\mathbf{q_{ij}}^\top\right) \\
&= \mu^{\mathrm{opt}}\mathrm{trace}\left(\mathbf{v}_2^{\mathrm{opt}}\mathbf{v}_2^{\mathrm{opt}\top}\sum_{(i,j)\in\mathcal{E}} w_{ij}^{\mathrm{opt}}\mathbf{q_{ij}}\mathbf{q_{ij}}^\top\right) \\
&= \mu^{\mathrm{opt}}\mathrm{trace}\left(\mathbf{v}_2^{\mathrm{opt}}\mathbf{v}_2^{\mathrm{opt}\top}L(w^{\mathrm{opt}})\right) \\
&= \mu^{\mathrm{opt}}\mathrm{trace}\left(\mathbf{v}_2^{\mathrm{opt}}\mathbf{v}_2^{\mathrm{opt}\top}\right) = \mu^{\mathrm{opt}}
\end{aligned}$$

where we use the fact that $\mathbf{v}_2^{\mathrm{opt}\top}L(w^{\mathrm{opt}}) = \mathbf{v}_2^{\mathrm{opt}\top}$, since $\lambda_2(w^{\mathrm{opt}}) = 1$. ∎

Theorems 1 and 2 with the iterations (11)-(12) provide a way to compute an optimal solution for the FMMP problem (4), as well as for the non-convex MVU problem (1). In fact, recalling that by equation (7), $x^{\mathrm{opt}} = \sqrt{c^{\mathrm{opt}}}\mathbf{v}_2^{\mathrm{opt}}$, once the couple $(\mu^{\mathrm{opt}}, \mathbf{v}_2^{\mathrm{opt}})$ is available through the iterations (11)-(12), due to the equivalence $\mu^{\mathrm{opt}} = c^{\mathrm{opt}}$, one can readily compute x^{opt} (in the case $\lambda_2(w^{\mathrm{opt}})$ is isolated).

Notice that the iterations (11)-(12) are not yet distributed, since they require the knowledge of the algebraic connectivity and its associated eigenvector.

V. Approximate Distributed Solution

In this section we propose a way to distribute the computation of the subgradients of $\mathcal{L}(w, \mu)$. Furthermore, we will analyze the case in which these subgradients are affected by some approximation error. This case is of practical importance when the communication effort among the nodes has to be limited, and therefore the iterative distributed algorithm to compute the subgradients of $\mathcal{L}(w, \mu)$ has to be stopped before reaching convergence.

Since $L(w)$ is a sparse matrix, we can now utilize the already mentioned distributed algorithm [9] to compute its eigenvalues and eigenvector. This technique, named by the authors as DOI algorithm (for decentralized orthogonal iteration), computes the spectral decomposition of a matrix M that exhibits the same sparsity of the underlying graph $\mathcal{G}$. The method is based on a QR decomposition and a consensus iteration and converges to within an accuracy of ε of the eigenspace of the matrix $M \in \mathbb{R}^{N \times N}$ in $O(\log^2(N/\varepsilon)1/\lambda_2(\mathcal{G}))$ rounds of communication/computations. Using this method, each node of the network has a copy of $\mathbf{v}_2$ with which they can compute locally the algebraic connectivity (by the multiplication of $\mathbf{v}_2$ with their row of the Laplacian).

The iterations (11)-(12) with the DOI algorithm could be used to solve the FMMP/MVU problems in a distributed way. It remains to prove that the convergence result of Theorem 1 still holds if the subgradients of the Lagrangian function (i.e., the algebraic connectivity and its associated eigenvector) are computed up to a prescribed accuracy ε.

Let $\nabla_{w,\varepsilon}\mathcal{L}(w, \mu)$ and $\nabla_{\mu,\varepsilon}\mathcal{L}(w, \mu)$ be ε-subgradients of $\mathcal{L}(w, \mu)$ with respect to w and μ, respectively. Consider the modification of the iterates (11)-(12) as

$$w^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+}\left[w^{(\tau)} - \alpha\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right] \quad (13)$$

$$\mu^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+}\left[\mu^{(\tau)} + \alpha\nabla_{\mu,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right] \quad (14)$$

Moreover, consider the modification of Assumption 1 as

Assumption 2: The ε-subgradients $\nabla_{w,\varepsilon}\mathcal{L}(w, \mu)$ and $\nabla_{\mu,\varepsilon}\mathcal{L}(w, \mu)$ used in the method defined by (13)-(14) are uniformly bounded, i.e., there is a constant $\Lambda_\varepsilon > 0$ such that

$$\max\left\{\left\|\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right\|, \left\|\nabla_{\mu,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right\|\right\} \leq \Lambda_\varepsilon$$

for all iterations $\tau \geq 0$.
We formalize the convergence of the running averages $\bar{w}^{(\tau)}$ and $\bar{\mu}^{(\tau)}$ based on the approximate iterates (13)-(14) to a saddle-point of $\mathcal{L}(w, \mu)$ in the following theorem.

Theorem 3: Under Assumption 2 the following relations for the iterates (13)-(14) hold true:

(a) For all $\tau \geq 1$,

$$-\frac{||\mu^{(0)} - \mu^{\text{opt}}||^2}{2\alpha\tau} - \frac{\alpha\bar{\Lambda}_\varepsilon^2}{2} \leq$$
$$\frac{1}{\tau}\sum_{j=0}^{\tau-1}\mathcal{L}(w^{(j)}, \mu^{(j)}) - \mathcal{L}(w^{\text{opt}}, \mu^{\text{opt}}) \leq$$
$$\frac{||w^{(0)} - w^{\text{opt}}||^2}{2\alpha\tau} + \frac{\alpha\bar{\Lambda}_\varepsilon^2}{2}$$

(b) The averages $\bar{w}^{(\tau)}$ and $\bar{\mu}^{(\tau)}$ satisfy the following relation for all $\tau \geq 1$:

$$-\frac{||\mu^{(0)} - \mu^{\text{opt}}||^2 + ||w^{(0)} - \bar{w}^{(\tau)}||^2}{2\alpha\tau} - \alpha\bar{\Lambda}_\varepsilon^2 \leq$$
$$\mathcal{L}(\bar{w}^{(\tau)}, \bar{\mu}^{(\tau)}) - \mathcal{L}(w^{\text{opt}}, \mu^{\text{opt}}) \leq$$
$$\frac{||w^{(0)} - w^{\text{opt}}||^2 + ||\mu^{(0)} - \bar{\mu}^{(\tau)}||^2}{2\alpha\tau} + \alpha\bar{\Lambda}_\varepsilon^2$$

with $\bar{\Lambda}_\varepsilon^2 = \Lambda_\varepsilon^2 + 2\varepsilon/\alpha$.
Proof: The proof follows from Proposition 1 of [12]. First we prove that:

$$||w^{(\tau+1)} - w||^2 \leq ||w^{(\tau)} - w||^2 - 2\alpha(\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w, \mu^{(\tau)})) + \alpha^2\bar{\Lambda}_\varepsilon^2 \quad (15)$$

$$||\mu^{(\tau+1)} - \mu||^2 \leq ||\mu^{(\tau)} - \mu||^2 + 2\alpha(\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w^{(\tau)}, \mu)) + \alpha^2\bar{\Lambda}_\varepsilon^2 \quad (16)$$

In order to show (15), we can expand $||w^{(\tau+1)} - w||^2$ into

$$||w^{(\tau+1)} - w||^2 =$$
$$\left\|\mathcal{P}_{\mathbb{R}_+}\left[w^{(\tau)} - \alpha\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right] - w\right\|^2 \leq$$
$$||w^{(\tau)} - w||^2 - 2\alpha\left\langle\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), w^{(\tau)} - w\right\rangle + \Lambda_\varepsilon^2$$

where we use the non-expansion property of the projection. Based on the definition of ε-subgradient and due to the convexity of the function $\mathcal{L}(w, \mu)$ in w

$$\left\langle\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), w - w^{(\tau)}\right\rangle - \varepsilon \leq$$
$$\mathcal{L}(w, \mu^{(\tau)}) - \mathcal{L}(w^{(\tau)}, \mu^{(\tau)})$$

$$-\left\langle\nabla_{w,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), w^{(\tau)} - w\right\rangle \leq$$
$$-(\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w, \mu^{(\tau)})) + \varepsilon$$

and therefore

$$||w^{(\tau+1)} - w||^2 \leq ||w^{(\tau)} - w||^2 - 2\alpha(\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w, \mu^{(\tau)})) + \alpha^2\Lambda_\varepsilon^2 + 2\alpha\varepsilon$$

which is (15).
In order to show (16), we can expand $||\mu^{(\tau+1)} - \mu||^2$ into

$$||\mu^{(\tau+1)} - \mu||^2 =$$
$$\left\|\mathcal{P}_{\mathbb{R}_+}\left[\mu^{(\tau)} + \alpha\nabla_{\mu,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)})\right] - \mu\right\|^2 \leq$$
$$||\mu^{(\tau)} - \mu||^2 + 2\alpha\left\langle\nabla_{\mu,\varepsilon}\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), \mu^{(\tau)} - \mu\right\rangle + \Lambda_\varepsilon^2$$

where we use the non-expansion property of the projection. Based on the definition of ε-subgradient and due to the concavity of the function $\mathcal{L}(w, \mu)$ in μ

$$\left\langle \nabla_{\mu, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), \mu - \mu^{(\tau)} \right\rangle + \varepsilon \geq \\ \mathcal{L}(w^{(\tau)}, \mu) - \mathcal{L}(w^{(\tau)}, \mu^{(\tau)})$$

$$\left\langle \nabla_{w, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)}), \mu^{(\tau)} - \mu \right\rangle \leq \\ (\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w^{(\tau)}, \mu)) + \varepsilon$$

and therefore

$$\|\mu^{(\tau+1)} - \mu\|^2 \leq \|\mu^{(\tau)} - \mu\|^2 + \\ 2\alpha(\mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) - \mathcal{L}(w^{(\tau)}, \mu)) + \alpha^2 \Lambda_\varepsilon^2 + 2\alpha\varepsilon$$

which is (16).

The proof of Proposition 1 of [12], i.e., Theorem 1, is based on (15) and (16) with $\varepsilon = 0$. Substituting Λ^2 with $\bar{\Lambda}_\varepsilon^2 = \Lambda_\varepsilon^2 + 2\varepsilon/\alpha$ it is not difficult to see that the analysis in [12] still holds and therefore the claim. ∎

Theorems 3 and 2 with the iterations (13)-(14) provide a way to compute an approximately optimal solution for the FMMP problem (4), as well as for the non-convex MVU problem (1) in a distributed way as summarized in Algorithm 1.

Algorithm 1 Primal-Dual Algorithm for the MVU and FMMP problems

1: Input $w^{(\tau)}, \mu^{(\tau)}$
   ▷ `Available data:` $\alpha, \varepsilon, L(w), (r_{ij}|\text{for each}(i, j) \in \mathcal{E})$
2: Determine $\lambda_2(w)$ and $\mathbf{v}_2$ of $L(w^{(\tau)})$ via the distributed DOI algorithm of [9] up to an accuracy of ε
3: Compute: $\nabla_{w, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)})$ and $\nabla_{\mu, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)})$ as

$$\begin{aligned} (\nabla_{w, \varepsilon} \mathcal{L}(w, \mu))_{ij} &= r_{ij}^2 - \mu(\mathbf{v}_{2i} - \mathbf{v}_{2j})^2 \\ \nabla_{\mu, \varepsilon} \mathcal{L}(w, \mu) &= 1 - \lambda_2(w) \end{aligned}$$

4: Compute:

$$\begin{aligned} w^{(\tau+1)} &= \mathcal{P}_{\mathbb{R}_+} \left[w^{(\tau)} - \alpha \nabla_{w, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) \right] \\ \mu^{(\tau+1)} &= \mathcal{P}_{\mathbb{R}_+} \left[\mu^{(\tau)} + \alpha \nabla_{\mu, \varepsilon} \mathcal{L}(w^{(\tau)}, \mu^{(\tau)}) \right] \end{aligned}$$

5: Compute the iteration of the MVU problem: $x^{(\tau)} = \sqrt{\mu^{(\tau)}} \mathbf{v}_2$
6: Output: $w^{(\tau+1)}, \mu^{(\tau+1)}, x^{(\tau)}$

In the next section we show a numerical example to assess the performance of the proposed algorithm.

VI. Numerical Example

We use a numerical example from [2] to show the performance of the primal-dual iterations (13)-(14) with the DOI algorithm applied to the FMMP and MVU problems, i.e., (4) and (1). Let the edge-distance constraints be specified as r_{ij} be $r_{12} = 1$, $r_{13} = 2$, $r_{23} = 1$, $r_{34} = 1$, $r_{45} = 1$, and $r_{46} = 2$. Figure 1 gives a pictorial representation of the problem.

The unique primal optimal solution (up to a multiplication by an orthogonal matrix) for the MVU problem (1) is

$$x^{\text{opt}} = (-2.5, -1.5, -0.5, 0.5, 1.5, 2.5)^\top$$
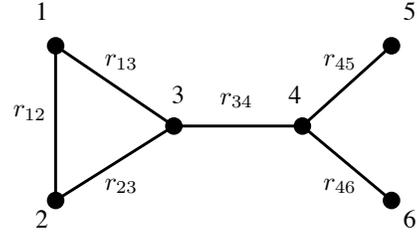
Fig. 1. Representation of the graph of the numerical example.
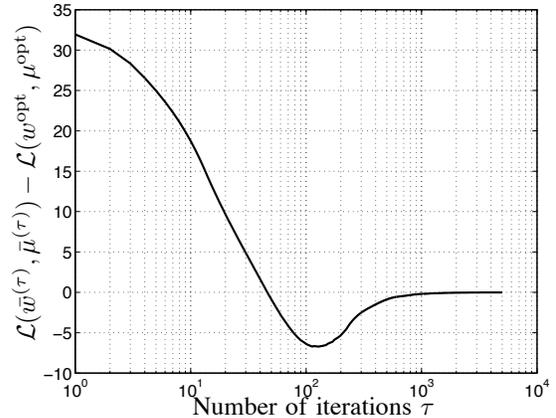
Fig. 2. Convergence of $\mathcal{L}(\bar{w}^{(\tau)}, \bar{\mu}^{(\tau)})$ to a saddle point of $\mathcal{L}(w, \mu)$ with respect to the iteration number τ.

while the optimal set of Lagrangian multipliers w_{ij} (thus the set of solutions of the dual FMMP problem (4)) is:

$$w_{12}^{\text{opt}} = u, \quad w_{23}^{\text{opt}} = 1.5 + u, \quad w_{13}^{\text{opt}} = 1.25 - 0.5u,$$

$$w_{34}^{\text{opt}} = 4.5, \quad w_{45}^{\text{opt}} = 1.5, \quad w_{46}^{\text{opt}} = 1.25$$

with the parameter $u \in [0, 2.5]$. The achieved optimal cost is $c^{\text{opt}} = 17.5$.

We use the iterations (13)-(14) with $\alpha = 0.15$ and $\varepsilon = 0.01$, starting from a random initial condition for $w^{(0)}$ and $\mu^{(0)}$. Figure 2 illustrates the convergence of $\mathcal{L}(\bar{w}^{(\tau)}, \bar{\mu}^{(\tau)})$ to a saddle point of $\mathcal{L}(w, \mu)$. We report that the total number of communication and computation iterations of the DOI algorithm, per iteration τ, was on average 12, and the computations required around 0.4 ms per node per iteration τ, on an Intel Core *i5* (2.3 GHz and 4GB DDR3) laptop. This leads to a total required time of 2 s if the scheme is run up to $\tau = 5000$. These communication/computation requirements are considered acceptable in many applications, especially in sensor network scenarios.

Finally we report the achieved tolerances

$$|\mathcal{L}(\bar{w}^{(5000)}, \bar{\mu}^{(5000)}) - \mathcal{L}(w^{\text{opt}}, \mu^{\text{opt}})| = 0.01,$$

$$|\lambda_2(\bar{w}^{(5000)}) - 1| = 0.03$$

which are acceptable given the value of ε and α.

VII. Extension to Multi-dimensional Problems and Localization Applications

In this section we extend the previous results to the case in which the variable x_i is a vector, meaning $x_i \in \mathbb{R}^n$, with $n > 1$. This scenario is particularly useful in localization problems [1]. In order to see this, consider N different sensor nodes sparsely placed in an n dimensional space. Let $x_i \in \mathbb{R}^n$ be the position of sensor node i. Assume that each node can determine its distance to the closest neighboring nodes and let r_{ij} be this distance for each connected couple of nodes (i, j). The problem of determining all the positions of the nodes via the measurements r_{ij} is called (anchor-free) localization problem and it can be written as [1]

$$\underset{x_i,\ldots,x_N}{\text{maximize}} \quad \sum_{i=1}^{N} ||x_i||^2 \tag{17a}$$

$$\text{subject to} \quad ||x_i - x_j||^2 = r_{ij}^2 \quad \forall (i,j) \in \mathcal{E} \tag{17b}$$

$$\sum_{i=1}^{N} x_i = 0 \tag{17c}$$

Problem (17) is similar to the MVU problem (1) with equality constraints instead of inequalities, and via minor modifications [2] one can translate one problem to the other.

In order to solve the MVU problem (1) for a multi-dimensional case, we proceed in the same way as in the scalar scenario. First, we define the matrix $X \in \mathbb{R}^{N \times N}$ as the Gramian matrix $X = xx^\top$, where $x \in \mathbb{R}^{Nn}$, and $x = (x_1, \ldots, x_N)^\top$. Then, we formulate the convex problem (2) (which is now a rank n relaxation) and its dual the FMMP problem (4). We note that these problems are not affected by x_i not being a scalar. Therefore all the analysis of convergence of Algorithm 1 is still valid for $x_i \in \mathbb{R}^n$. The only notable difference is that for the multi-dimensional case the geometric multiplicity of $\lambda_2(w)$ needs[3] to be greater than or equal to n [2]. This implies that at least $\lambda_2 = \lambda_3 = \cdots = \lambda_{n+1} = 1$ and therefore, the optimal X can be written as

$$X^{\text{opt}} = \frac{c^{\text{opt}}}{n} \sum_{i=2}^{n+1} \mathbf{v}_i^{\text{opt}} \mathbf{v}_i^{\text{opt}\top} \tag{18}$$

while

$$x = \sqrt{\frac{c^{\text{opt}}}{n}} (\mathbf{v}_2^{\text{opt}}, \ldots, \mathbf{v}_{n+1}^{\text{opt}})^\top \tag{19}$$

With these relations we can compute the optimal value for x_i and solve the multi-dimensional MVU problem (1) in a distributed way via Algorithm 1.

VIII. A Dispersion Problem for Mobile Networks

In Section VI we have considered a numerical example where the nodes were fixed. This type of problem is typical in sensor network applications where the network structure is predetermined. Interesting application perspectives arise when

---

[3]This is related to the rank relaxation and it may seems a restrictive condition to fulfill, however it works rather well in practice [1]. Further analysis of this aspect are ongoing research topics.

we let the nodes move as in [10], and the weights depend on the position of the nodes.

Let $\mathbf{x}_i \in \mathbb{R}^2$ be the position of the mobile node i. Let $\mathbf{x}$ be the stacked vector of all the node positions. We let the weights w_{ij} depend on the positions as

$$w_{ij}(\mathbf{x}) = f_w(||\mathbf{x}_i - \mathbf{x}_j||^2)$$

where $f_w : \mathbb{R} \to [0, 1]$ is a smooth function of the squared Euclidean distance. As in [13], f_w will be 0 for distances greater than a certain value and 1 for distances smaller than a given threshold. Within these values f_w varies smoothly.

We write $w_{ij}(\mathbf{x})$ and $\lambda_2(\mathbf{x})$ to highlight the dependence of the weights and of the algebraic connectivity on the nodes' positions.

The FMMP problem (4) can then be rewritten as the following non-convex optimization problem

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{(i,j) \in \mathcal{E}} r_{ij}^2 w_{ij}(\mathbf{x}) \tag{20a}$$

$$\text{subject to} \quad \lambda_2(\mathbf{x}) > \bar{\lambda}_2 \tag{20b}$$

where the decision variables are the node locations $\mathbf{x}$, while $\bar{\lambda}_2$ is a prescribed level of connectivity.

Problem (20) can be seen as the maximization of the dispersion (i.e., the distance among the nodes) with the guarantee of the maintenance of a prescribed level of connectivity. This problem has been studied in the robotics literature [14]–[19], however in the mentioned works no clear guarantees to obtain a prescribed level of connectivity are given.

Problem (20) can be solved by a two-step procedure, similarly to [10] as follows.

1) Start with a given $\mathbf{x}^{(0)}$ and $\mu^{(0)}$. Set $\tau = 0$.
2) Compute $w^{(\tau)}(\mathbf{x}^{(\tau)})$ using f_w.
3) Determine the vector iterates $w^{(\tau+1)}, \mu^{(\tau+1)}$ with the distributed Algorithm 1, using

$$w^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+} \left[w^{(\tau)}(\mathbf{x}^{(\tau)}) - \alpha \nabla_{w,\varepsilon} \mathcal{L}(w^{(\tau)}(\mathbf{x}^{(\tau)}), \mu^{(\tau)}) \right]$$

$$\mu^{(\tau+1)} = \mathcal{P}_{\mathbb{R}_+} \left[\mu^{(\tau)} + \alpha \nabla_{\mu,\varepsilon} \mathcal{L}(w^{(\tau)}(\mathbf{x}^{(\tau)}), \mu^{(\tau)}) \right]$$

This is equivalent to solving the convex FMMP problem (4) using only one iteration.

4) Compute the position $\mathbf{x}^{(\tau+1)}$ using a potential field approach, just as in [10]. This step can be seen as finding the $\mathbf{x}^{(\tau+1)}$ that solves the non-convex problem

$$\underset{\mathbf{x}^{(\tau+1)}}{\text{minimize}} \quad ||w^{(\tau+1)} - w^{(\tau+1)}(\mathbf{x}^{(\tau+1)})||^2$$

5) Set $\tau \leftarrow \tau + 1$ and go back to step 2).

Under the same assumptions as Theorem 3, this two step-procedure will converge to the saddle point of the convex FMMP problem (4) within the error level $\alpha \bar{\Lambda}_\varepsilon^2$, where ε, in this case, incorporates also the errors in the computation of $w^{(\tau+1)}$ due to the potential field approach.

In Figures 3 and 4 we represent a numerical example consisting of 9 mobile nodes that are moved in a 2D plane to obtain an optimal dispersion given a bound on the algebraic connectivity, $\bar{\lambda}_2 \geq 3/4$.
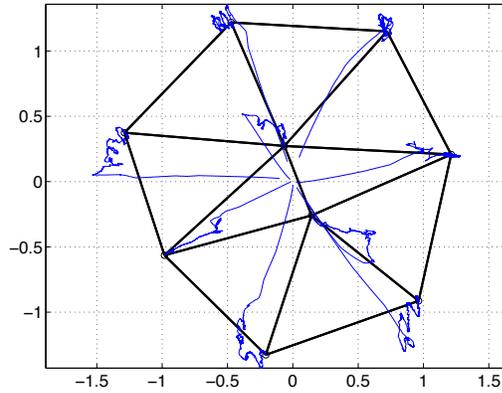
Fig. 3. Final graph in black and node position iterates as blue trajectories. The initial node positions are random and centered around the origin.
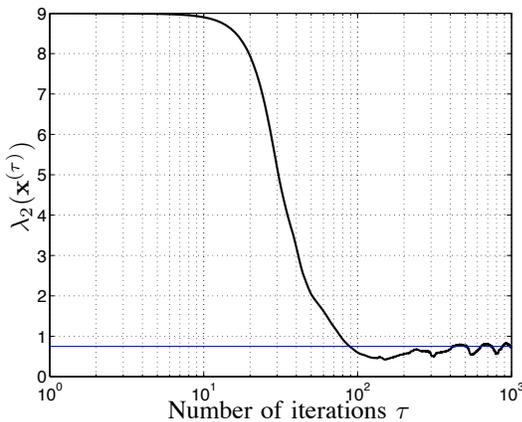
Fig. 4. Convergence of the algebraic connectivity to the desired level.

Figure 3 depicts the iterations on the node positions as blue trajectories, and the final configuration of the graph (with black lines). Figure 4 shows the convergence of the two-step procedure to the desired level of connectivity within an error level (as one would expect from the previous discussion).

IX. CONCLUSIONS

In this paper we have focused on the non-convex networked optimization problem called MVU and its dual formulation, the FMMP. We have solved them using a primal-dual approximate subgradient method and proved the convergence of the solution using a fixed step-size. We have illustrated the relevance of the algorithm in sensor network and mobile robotic network applications. Further research directions encompass acceleration of the convergence of Algorithm 1 via regularization techniques [20], [21] and robustness to measurement errors for the localization scenario (17).

REFERENCES

[1] K. Q. Weinberger, F. Sha, Q. Zhu, and L. K. Saul, *Advances in Neural Information Processing Systems 19*. Cambridge, MA: MIT Press, 2007, ch. Graph Laplacian Regularization for Large-Scale Semidefinite Programming, pp. 1489 – 1496.

[2] J. Sun, S. Boyd, L. Xiao, and P. Diaconis, "The Fastest Mixing Markov Process on a Graph and a Connection to a Maximum Variance Unfolding Problem," *SIAM Review*, vol. 48, no. 4, pp. 681 – 699, 2006.

[3] F. Göring, C. Helmberg, and M. Wappler, "Embedded in the Shadow of the Separator," *SIAM Journal on Optimization*, vol. 19, no. 1, pp. 472 – 501, 2008.

[4] K. Q. Weinberger and L. K. Saul, "Unsupervised Learning of Image Manifolds by Semidefinite Programming," *International Journal of Computer Vision*, vol. 70, no. 1, pp. 77 – 90, 2006.

[5] C. Godsil and G. Royle, *Algebraic Graph Theory*, ser. Graduate text in Mathematics. Springer, 2001.

[6] D. P. Bertsekas, *Constrained Optimization and Lagrangian Multiplier Methods*. Accademic Press, 1982.

[7] ——, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

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

[9] D. Kempe and F. McSherry, "A Decentralized Algorithm for Spectral Analaysis," *Journal of Computer and System Sciences*, vol. 74, no. 1, pp. 70 – 83, 2008.

[10] M. C. De Gennaro and A. Jadbabaie, "Decentralized Control of Connectivity for Multi-Agent Systems," in *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006, pp. 3628 – 3633.

[11] K. Kiwiel, "Convergence of Approximate and Incremental Subgradient Methods for Convex Optimization," *SIAM Journal of Optimization*, vol. 14, no. 3, pp. 807 – 840, 2004.

[12] A. Nedić and A. Ozdaglar, "Subgradient Methods for Saddle-Point Problems," *Journal of Optimization Theory and Applications*, vol. 142, no. 1, pp. 205 – 228, 2009.

[13] Y. Kim and M. Mesbahi, "On Maximizing the Second Smallest Eigenvalue of a State-Dependent Graph Laplacian," *IEEE Transactions of Automatic Control*, vol. 51, no. 1, pp. 116 – 120, 2006.

[14] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Mobile Sensor Network Deployment using Potential Fields: a Distributed, Scalable Solution to the Area Coverage Problem," in *Proceedings of the 6th International Symposium on Distributed an Autonomous Systems*, Fukuoka, Japan, June 2002, pp. 299 – 208.

[15] J. Cortés, S. Martinez, T. Karatas, and F. Bullo, "Coverage Control for Mobile Sensing Networks," *IEEE Transaction on Robotics and Automation*, vol. 20, no. 2, pp. 243 – 255, 2004.

[16] V. A. Susan and S. Dubowsky, "Visually Guided Cooperative Robot Actions Based on Information Quality," *Autonomous Robots*, vol. 19, pp. 89 – 110, 2005.

[17] A. Arsie and E. Frazzoli, "Efficient Routing of Multiple Vehicles with no Communications," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 2, pp. 154 – 164, 2007.

[18] I. I. Hussein and D. Stipanovic, "Effective Coverage Control for Mobile Sensor Networks with Guaranteed Collision Avoidance," *IEEE Transactions on Control System Technology*, vol. 15, no. 4, pp. 624 – 657, 2007.

[19] D. Dimarogonas and K. Kyriakopoulos, "Inverse Agreement Protocols With Application to Distributed Multi-Agent Dispersion," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 657 – 663, 2009.

[20] O. Devolder, F. Glineur, and Y. Nesterov, "A Double Smoothing Technique for Constrained Convex Optimization Problems and Applications to Optimal Control," *Submitted to SIAM Journal on Optimization*, 2011.

[21] J. Koshal, A. Nedić, and U. Y. Shanbhag, "Multiuser Optimization: Distributed Algorithms and Error Analysis," *SIAM Journal on Optimization*, vol. 21, no. 3, pp. 1046 – 1081, 2011.