2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

A Topological View of Estimation from Noisy Relative Measurements

Pooya Molavi and Ali Jadbabaie

Abstract— **In this paper we study the problem of estimating the state of sensors in a sensor network from noisy pairwise relative measurements. The underlying sensor network is typically modeled by a graph whose edges correspond to pairwise relative measurements and nodes represent sensors. Using tools from algebraic topology and cohomology theory, we present a new model in which the higher order relations between measurements are captured as simplicial complexes. This allows us to address the fundamental tension between two conflicting goals: finding estimates that are close to obtained measurements, and at the same time are *consistent* around any sequence of pairwise measurements that form a cycle. By defining a measure of inconsistency around each cycle, we present a one-parameter family of algorithms that solves the estimation problem by identifying and removing the smallest fraction of measurements that make the estimates globally inconsistent. We demonstrate that the inconsistencies are due to topological obstructions and can be decomposed into local and global components that have interesting geometric interpretations. Furthermore, we show that the proposed algorithm is naturally distributed and will provably result in consistent estimates, and more importantly, recovers two sparse estimation algorithms as special cases.**

I. Introduction

A common application of estimation in sensor networks is to find the best estimate of a quantity of interest using noisy relative measurements. Examples of problems of this type include localization and time synchronization where one is interested in determining some unknown absolute quantity represented by node values in a network using a set of pairwise differences. The unknown variable on the nodes can represent the time offset of a clock relative to a reference clock in the time synchronization problem or coordinates or direction of a moving sensor.

Nodes in a sensor network often have limited communication power. As a result, only a subset of all possible relative measurements is available to each sensor. The noisy measurements can be represented as edge valuations in a network and the unknown absolute quantities as node valuations. Simply put, the goal of such estimation is to use noisy relative measurements to find the absolute node values (cf. Karp *et al.* [1], and Giridhar *et al.* [2] on time synchronization in the sensor networks).

Since adding a constant to all the node values does not change the relative differences, the best we can hope for is to find an estimate of the pairwise differences rather than the absolute node values. Therefore, the problem is sometimes formulated as estimation of the pairwise differences from their noisy measurements. To make the resulting node values unique, it is assumed that there is an *anchor node* whose state is known and can be used as a reference for other nodes.

The two main objectives in estimation of pairwise differences are, first for the estimates to be close to the measured values, and second for the estimates to be *consistent*. A set of pairwise differences is consistent if it can be used (together with the state of the reference node) to define a unique estimate of the node values. For example, if there are three scalar pairwise measurements that form a clique in the underlying graph, one would like the estimated edge values for the 3-clique to have the additional property that the sum of two pairwise differences gives the third one, or in other words, borrowing from circuits terminology, Kirchoff's Voltage Law (KVL) is satisfied across every loop, exactly as in an electric circuit.

In [1] the authors provide a protocol for sensors to estimate the clock offsets in an optimal and globally consistent way. More recently, Barooah *et al.* in [3] study the effect of graph structure on estimation accuracy and provide lower and upper bounds on the variance of estimation error (MSE). They also look at vector valued variables and propose a decentralized and asynchronous algorithm to find the Best Linear Unbiased Estimator (BLUE) [4]. In [5], the authors suggest a new decentralized algorithm to find BLUE with fewer message exchanges. When the noise in relative measurements is Gaussian and independent, a BLUE can be computed using standard least squares by simply computing a pseudo-inverse [6]. On the other hand, to solve the problem using purely local information, an iterative coordinated descent algorithm can be used. Xiao *et al.* in [7] consider a general linear estimation model and use repeated local averaging to compute the least square solution even when the underlying network topology changes over time. When the linear measurement model involves pairwise relative differences, there is a beautiful analogy between solution of this least squares problem and electric network theory (cf. [2], and Doyle and Snell [8]).

While the above mentioned results and many others provide distributed (and sometimes asynchronous) solutions to the problem of finding a consistent edge valuation, the resulting estimates often suffer from the drawback that estimated edge values are different from the measured ones on most of the edges: large errors in a small subset of measurements can distort an otherwise consistent estimate on a large subset of edges. One could then ask the question: what are the fewest number of measurements that prevent the other measurements from forming a consistent estimate?

The authors are with the Department of Electrical and Systems Engineering and General Robotics, Automation, Sensing and Perception (GRASP) Laboratory, University of Pennsylvania, Philadelphia, PA 19104-6228 USA. {pooya,jadbabai}@seas.upenn.edu

This research is supported in parts by ONR MURI N000140810747 on Next Generation Network Science, AFOSR Complex Networks Program and DARPA DSO's SToMP program.

This is the approach taken in this paper. The answer to this question obviously depends on graph topology: On a graph without loops any edge valuation is consistent. On the other hand, large systematic errors in pairwise measurements cannot be detected on a cycle-free graph due to lack of redundancy. To answer the question posed, one needs to go beyond pairwise interactions and a graph-based model of the underlying sensor network and consider *simplicial complexes* as models of the network [9]–[11]. Our results are informed and motivated by the works in [9], [10], as well as Jiang *et al.* [12] in construction of absolute rankings from relative rankings.

In this paper we argue that there could be topological *obstructions* to aggregation of locally consistent edge valuations into a global estimate. We show that even in the presence of such topological obstacles, one can find the sparsest representation of *inconsistency-causing measurements*. We demonstrate how cohomology theory and its sparse representation can be used to identify the fewest number of measurements that need to be discarded to get a consistent estimate. We also explore the tradeoff between consistency of an estimate and how close to the measured values it is. We quantify the (local) inconsistency of an estimate and show through simulations how increasing the threshold of the inconsistency we are willing to accept in an estimate, makes the estimate closer to the measured value.

II. Problem Setup and Mathematical Preliminaries

Let $G = (V, E)$ be an undirected graph representing the network of sensors and relative measurements where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes and E is the set of edges with $m = |E|$. Two nodes v_i, v_j are said to be neighbors, denoted by $v_i \sim v_j$, when $\{v_i, v_j\} \in E$. Let $x : V \to \mathbb{R}$ be an unknown function that assigns a real value to each of the nodes. x is only known through measurements of pairwise differences of its value on neighboring nodes. Any two neighboring sensors v_i and v_j communicate with each other to form a *common measurement* of the difference between $x(v_i)$ and $x(v_j)$. As a consequence, even though noise is oftentimes at the node level, we can effectively attribute it to relative measurements. Noisy pairwise measurements are modeled as a real valued function y defined for all (v_i, v_j) such that $\{v_i, v_j\} \in E$ as

$$y(v_i, v_j) = x(v_i) - x(v_j) + \varepsilon(v_i, v_j), \quad (1)$$

where $\varepsilon(v_i, v_j) = -\varepsilon(v_j, v_i)$ is the noise of measurement. Given y, we want to find an estimate $\hat{x}$ of x. To measure the quality of $\hat{x}$, we compare y with $\hat{y}$, the pairwise differences that are induced on the edges, if $\hat{x}$ was the unknown function. $\hat{y}$ is related to $\hat{x}$ by $\hat{y}(v_i, v_j) = \hat{x}(v_i) - \hat{x}(v_j)$. Therefore, the problem can be reformulated as minimizing $\|y - \hat{y}\|$ subject to $\hat{y} : E \to \mathbb{R}$ being induced by pairwise differences of some function $\hat{x} : V \to \mathbb{R}$. Once $\hat{y}$ is known, $\hat{x}$ can be determined up to an additive constant. To get a unique $\hat{x}$, we assume that for every connected component of the graph there exist an anchor node on which the value of x is known.

We now present a brief review of the key concepts from algebraic topology that we use in this paper, namely simplicial complexes, cochains, and cohomology. The following definitions and results are standard and can be found in [13], [14].

A k dimensional *simplex* henceforth called a k-simplex is a set of $k+1$ points $\{v_0, v_1, \ldots, v_k\}$. For any $j \leq k$ a *j-face* of a k-simplex is a nonempty subset of the $k+1$ points that define the simplex.

Definition 1: A *simplicial complex* $\mathcal{K}$ is a finite collection of simplices that has two properties:

1) Every face of $\mathcal{K}$ is in $\mathcal{K}$.
2) The intersection of any two simplices in $\mathcal{K}$ is a face of each of them.

We can define an orientation for a simplicial complex by defining an ordering on all of its simplices. We denote the k-simplex $\{v_0, v_1, \ldots, v_k\}$ with some ordering on it by $[v_0, v_1, \ldots, v_k]$. Likewise, an orientation of a simplicial complex implies an ordering on all of its simplices.

The dimension of a simplicial complex is the maximum of the dimensions of its simplices. A *j-skeleton* of $\mathcal{K}$ is a simplicial complex consisting of all of the j-simplices of $\mathcal{K}$. In particular the 1-skeleton of a simplicial complex is a graph. For $k \geq 2$ the k dimensional *flag complex* of a graph G denoted by $F_k(G)$ is the largest simplicial complex of dimension k whose 1-skeleton is G.

Definition 2: Given an oriented simplicial complex $\mathcal{K}$, let C_k be the set of all its ordered k-simplices. A k-dimensional *cochain* over the field of reals henceforth called a k-cochain is a function $f : C_k \to \mathbb{R}$ such that

$$f([v_0, v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k]) = -f([v_0, v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k]),$$

for all $1 \leq i < j \leq k$.

The set of all k-cochains of a simplicial complex is a vector field over the reals denoted by $C^k(\mathcal{K})$. We let $C^k(\mathcal{K}) = 0$, if k is greater than the dimension of $\mathcal{K}$. Note that $C^k(\mathcal{K})$ is a finite dimensional vector field with its dimension equal to the number of the k-simplices in $\mathcal{K}$. Next we define coboundary maps.

Definition 3: The kth *coboundary operator* is a linear map $\delta_k : C^k(\mathcal{K}) \to C^{k+1}(\mathcal{K})$ defined as

$$(\delta_k f)([v_0, v_1, \ldots, v_{k+1}]) = \sum_{j=0}^{k+1} (-1)^j f([v_0, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{k+1}]).$$

It is easy to prove:

Lemma 1: The linear map $\delta_k \circ \delta_{k-1} : C^{k-1}(\mathcal{K}) \to C^{k+1}(\mathcal{K})$ is the zero map for all $k \geq 1$.

Given an oriented simplicial complex $\mathcal{K}$, for any $\sigma \in C_k$ whose ordering is consistent with the orientation of $\mathcal{K}$, let $b_\sigma : C_k \to \mathbb{R}$ be the k-cochain such that $b_\sigma(\sigma) = 1$ and $b_\sigma(\tau) = 0$ for all $\tau \in C_k$ which are not permutations of σ. The assumption that b_σ is a k-cochain uniquely specifies its value on ordered k-simplices which are permutations of σ. The set of all b_σ is the canonical basis for the vector space $C^k(\mathcal{K})$. The representation of a cochain with respect to a basis can be written as a column vector. Likewise the linear maps on cochains can be represented as matrices. We do not distinguish between cochains and coboundary maps, and their vector and matrix representations with respect to the canonical basis.

To define simplicial cohomology, we need to look at two subspaces of $C^k(\mathcal{K})$:

$$\ker\ \delta_k = \{f \in C^k(\mathcal{K}) : \delta_k f = 0\},$$
$$\text{im}\ \delta_{k-1} = \{f \in C^k(\mathcal{K}) : \exists\, g \in C^{k-1}(\mathcal{K}) \text{ s.t. } f = \delta_{k-1} g\}.$$

k-*cocylces* are elements of kernel of δ_k and k-*coboundaries* are elements of image of δ_{k-1}. By Lemma 1, im δ_{k-1} is a subspace of ker δ_k. Hence, we can define the kth *cohomology* $H^k(\mathcal{K})$ as the quotient vector space

$$H^k(\mathcal{K}) = \ker\ \delta_k / \text{im}\ \delta_{k-1}.$$

The dimension of kth cohomology is called the kth Betti number $\beta_k(\mathcal{K})$[1]. Informally, the kth Betti number counts the number of k dimensional 'holes' in the simplicial complex. For example, $\beta_0(\mathcal{K})$ is the number of connected components in $\mathcal{K}$, or $\beta_1(\mathcal{K})$ is the number of two-dimensional or 'circular' holes. If the simplicial complex is contractible (i.e. can be contracted continuously into a point), then all Betti numbers are equal to zero. Two cochains are said to be *cohomologous* if their difference is a coboundary.

Definition 4: The cohomology class of the cochain $f \in C^k(\mathcal{K})$ denoted by $[f]$ is the set

$$[f] = \{g \in C^k(\mathcal{K}) : g - f \in \text{im}\ \delta_{k-1}\}.$$

Hodge decomposition is an orthogonal decomposition of the vector space of cochains into three subspaces. To state the Hodge decomposition theorem, we first need to define the adjoint operators of coboundary maps. Let $\langle\cdot,\cdot\rangle_k$ and $\langle\cdot,\cdot\rangle_{k+1}$ be inner products on $C^k(\mathcal{K})$ and $C^{k+1}(\mathcal{K})$ respectively. Then, the unique adjoint operator corresponding to the kth coboundary operator is $\delta_k^* : C^{k+1}(\mathcal{K}) \to C^k(\mathcal{K})$ defined to satisfy $\langle \delta_k f, g\rangle_{k+1} = \langle f, \delta_k^* g\rangle_k$, for all $f \in C^k(\mathcal{K})$ and $g \in C^{k+1}(\mathcal{K})$. Unless otherwise specified, all the inner products are the standard inner products of vector representations of cochains with respect to the canonical basis.

Definition 5: The kth combinatorial Laplacian operator is the linear operator $\Delta_k : C^k(\mathcal{K}) \to C^k(\mathcal{K})$ defined as

$$\Delta_k = \delta_{k-1} \circ \delta_{k-1}^* + \delta_k^* \circ \delta_k.$$

We now can state the Hodge decomposition theorem. A simple proof can be found in [15].

Theorem 1 (Hodge decomposition): Vector space of k-cochains can be decomposed orthogonally as

$$C^k(\mathcal{K}) = \text{im}\,\delta_{k-1} \oplus \ker \Delta_k \oplus \text{im}\,\delta_k^*. \quad (2)$$

Furthermore, the dimension of $\ker \Delta_k$ equals $\beta_k(\mathcal{K})$.

Therefore, $\ker \Delta_k$ is trivial if and only if $\beta_k(\mathcal{K}) = 0$.

[1]Betti numbers are really defined as dimensions of homology and not cohomology groups. However, since our coefficient group is the field of reals, universal coefficient theorem for cohomology implies that the two are the same (see Theorem 53.5 in [14]).

III. Problem Reformulation on Flag Complexes

In this section we reformulate the estimation problem as finding a projection of the 1-cochain of measurements onto the consistent subspace of the flag complex of the graph. Let $G = (V, E)$ be the graph defined in section I, and let $\mathcal{K} = F_2(G)$ be its two dimensional flag complex. Let x, y, and ε be defined as in section I. It is easy to verify that x is a 0-cochain whereas y and ε are 1-cochains. Equation (1) can be rewritten in terms of coboundary operator as $y = \delta_0 x + \varepsilon$.

The estimation problem as presented in section II is an optimization problem with the constraint that $\hat{y}$ is induced by pairwise differences of some $\hat{x} : V \to \mathbb{R}$. We can reformulate the constraint using the terminology of algebraic topology as: $\hat{y}$ is a 1-cochain that lies in im δ_0. This is the subspace of 'consistent' 1-cochains, namely the cochains that can be written as the coboundary of some 0-cochain. The name consistent refers to the fact that 1-cochains in im δ_0 can give rise to a node valuation which is unique up to an additive constant. The problem of finding a consistent edge valuation from noisy pairwise measurements can be written as:

$$\begin{aligned} \min_{\hat{y} \in C^1(\mathcal{K})} \quad & \|y - \hat{y}\| \\ \text{subject to} \quad & \hat{y} \in \text{im}\,\delta_0, \end{aligned} \quad (3)$$

where different norms result in different notions of 'best' estimate. In what follows we look at the solution and the residual of (3), when the norm that is used is the ℓ^2 norm. We use Hodge theory to decompose the residual and motivate our estimation algorithm.

A. Orthogonal Projection

A commonly used norm in (3) is the ℓ^2 norm: $\|f\|_2 = \langle f, f\rangle_1^{1/2}$. Let the inner product on $C^1(\mathcal{K})$ be defined as $\langle f, g\rangle_1 = f^T W g$, where W is some positive semidefinite matrix. Then the solution $\hat{y}^*$ to the optimization problem (3) is just the orthogonal projection of y, with respect to $\langle\cdot,\cdot\rangle_1$, onto im δ_0. $\hat{y}^*$ is the consistent component of y, whereas the residual $r^* \triangleq y - \hat{y}^*$ corresponds to the inconsistencies in y. Hodge decomposition theorem implies that the residual can be further decomposed into the sum of two orthogonal parts such that one captures the inconsistencies in y on a local level whereas the other captures it on a global level. This is the approach taken in [12] in the context of ranking. We use the same terminology as [12] to name the components of a 1-cochain in its Hodge decomposition as 'consistent', 'locally inconsistent', and 'globally inconsistent'.

Image of δ_0 is the subspace of 1-cochains that can arise as the pairwise differences of a function on nodes of the graph. Hence, we call it the consistent component. Image of δ_1^* is the space of 1-cochains that sum up to a non-zero value on the boundary of some 2-simplex. More precisely, for any 1-cochain $f \in$ im δ_1^*, there exists an unordered 2-simplex $\{v_1, v_2, v_3\}$ in $\mathcal{K}$, such that $f([v_1, v_2]) + f([v_2, v_3]) + f([v_3, v_1]) \neq 0$. We call a 1-cochain in im δ_1^* locally inconsistent, because even on nodes of the 2-simplex (i.e. locally), it cannot give rise, in a consistent way, to a node valuation. Finally, $\ker \Delta_1$ is the *harmonic* or globally inconsistent (but locally consistent) subspace of $C^1(\mathcal{K})$. This is the space of cochains that sum up to zero on all 2-simplices but cannot be written consistently as differences of node valuations.

Remark 1: One can think of a cochain in im δ_0 as the gradient of some scalar potential defined on the nodes of the graph. Similarly, a cochain in im δ_1^* can be thought of as being the curl of a vector field on 2-simplices. A well-known result from vector calculus asserts that a curl free vector field can be written as the gradient of some scalar potential, i.e. the space of 1-cochains can be decomposed as $\mathrm{im}\,\delta_0 \oplus \mathrm{im}\,\delta_1^*$. This result is only true for a space with a zero β_1 (e.g. $\mathbb{R}^n$). In general a 1-cochain might also have a component in $\ker \Delta_1$.

B. Sparse Cohomologous Cochains

The alternative algorithm we peruse here is to find the fewest problematic measurements and discard them to make the remaining ones consistent. Even though a small fraction of the measurements are ignored altogether in our algorithm, but the remaining ones are not being altered.

Hodge decomposition theorem implies that if $\ker \Delta_1$ is trivial, then any 1-cochain that is locally consistent is also a globally consistent solution to the estimation problem. A sufficient condition for $\beta_1(\mathcal{K}) = 0$ is for $F_2(G)$, the two dimensional flag complex of G, to be the disjoint union of finitely many simply connected sets.

When im δ_1^* is trivial, we can look at the projection of y onto $\ker \Delta_1$ to find out the 1-cochains that prevent the locally consistent 1-cochains from forming a globally consistent estimate. Otherwise, we first orthogonally project y onto $\mathrm{im}\,\delta_0 \oplus \ker \Delta_1$ and then proceed as before. We will provide an alternative solution to deal with the local inconsistencies in the next section.

The 1-cochain resulting after projecting out the locally inconsistent component is the sum of a locally consistent component and a harmonic component which is the *topological obstruction* to the aggregation of the locally consistent component into a globally consistent estimate. This obstruction is captured by the cohomology class of the cochain. $[f]$ is the set of all 1-cochains that differ from f in a consistent 1-cochain. All the 1-cochains in the same cohomology class represent the same inconsistency. In Lemma 2 we prove that if we remove the support of any $g \in [f]$ from the measurement graph, f becomes a consistent 1-cochain on the remaining edges of the graph. Therefore, given a locally consistent 1-cochain f, we can get a globally consistent one by removing the edges in the support of any cohomologous cochain g. Since we would like to discard the fewest number of edges possible, we are interested in finding the sparsest cochain which is cohomologous to the harmonic component.

For any k and $f \in C^k(\mathcal{K})$, the k-cochain in $[f]$ with the smallest support can be found by solving the following non-convex optimization problem:

$$\min_{g \in C^{k-1}(\mathcal{K})} \quad \|f + \delta_{k-1} g\|_0, \tag{4}$$

where $\|\cdot\|_0$, sometimes called the zero norm, is defined as $\|f\|_0 = \mathrm{card}(\{(v_i, v_j) : f([v_i, v_j]) \neq 0\})$. A commonly used convex relaxation of a zero norm minimization problem is to replace the zero norm with the ℓ^1 norm [10]. This is the approach we take here.

Having motivated all the steps, we now present our algorithm for finding an estimate $\hat{x}^*$ by removing the fewest number of problematic edges. For the estimate to be unique, we assume that there is an anchor node on each connected component of the sensor network.

1) Project y onto $\mathrm{im}\,\delta_0 \oplus \ker \Delta_1$ orthogonally to get $\hat{y}^*$, the locally consistent component of y.
2) Project y onto $\ker \Delta_1$ orthogonally to get h^*, the harmonic component of $\hat{y}^*$.
3) Solve the following optimization problem to find g^*:
$$\min_{g \in C^0(\mathcal{K})} \quad \|h^* + \delta_0 g\|_1. \tag{5}$$
$h^* + \delta_0 g^*$ is a 1-cochain with a small support which is cohomologous to h^*.
4) Remove the edges in the support of $h^* + \delta_0 g^*$ from the measurement graph.
5) Use $\hat{y}^*$ on the edges not deleted to uniquely find the estimate $\hat{x}^*$.

To prove that this algorithm uniquely defines an estimate $\hat{x}^*$, we need to show that: (a) $\hat{y}^*$ becomes globally consistent after step 4 of the algorithm and (b) $\hat{y}^*$ uniquely defines an estimate $\hat{x}^*$ on the remaining edges.

Lemma 2: After step 4 of the algorithm, $\hat{y}^*$ is globally consistent on the remaining edges.

Proof: Let $y = y_g + y_h + y_c$ where $y_g \in$ im δ_0, $y_h \in$ ker Δ_1, and $y_c \in$ im δ_1^*. $\hat{y}^* = y_g + y_h$ and $h^* = y_h$ imply that $h^* + \delta_0 g^* = y_h + \tilde{y}_g$ for some $\tilde{y}_g \in$ im δ_0. In step 4 of the algorithm we only delete edges on which $h^* + \delta_0 g^* \neq 0$. Hence, on all the remaining edges we have $y_h = -\tilde{y}_g$. Since $\tilde{y}_g$ is the coboundary of some 0-cochain which is not affected by edge deletion, $\tilde{y}_g$ will still be in im δ_0 after step 4. Therefore, on all the remaining edges $\hat{y}^* = y_g - \tilde{y}_g$ for some $\tilde{y}_g \in$ im δ_0. ■

We also show in Lemma 4 that the connected components of the graph do not become disconnected in the edge deletion step. Therefore, there exists a path connecting each node v_i to some anchor node v_0. Consequently, $\hat{x}^*(v_i)$ can be found by adding $x(v_0)$ to the sum of $\hat{y}^*$ on the path connecting the two. Consistency of $\hat{y}^*$ implies that $\hat{x}^*$ is independent of the path chosen. To prove that the connectivity of the graph does not change after running the algorithm, we first state a lemma.

Lemma 3: For any optimal solution a^* to the problem

$$\min_{a} \quad \sum_{k=1}^{p} |z_k - a|, \tag{6}$$

which is an extreme point, $z_k = a^*$ for some $k \in \{1, \ldots, p\}$.

The proof is trivial and is omitted. For more about extreme points and linear optimization see [16].

Lemma 4: If $h^* + \delta_0 g^*$ is an extreme point, connected components stay connected after the edge deletion step of the algorithm.

Proof: Without loss of generality we can assume that the graph has only one connected component to start with. We prove the lemma by contradiction. Assume that after deleting the edges we end up with at least two components. Then there exists at least an edge $\{v_i, v_j\} \in E$ such that v_i is in component 1 and v_j is in the component 2 and $(h^* + \delta_0 g^*)([v_i, v_j]) \neq 0$. Let $\tilde{E}$ be the set of such edges and let $p = |\tilde{E}|$. Assume that there exists one $\{v_i, v_j\} \in \tilde{E}$ such that $\hat{y}^*([v_i, v_j]) \neq 0$. This assumption is generic and is satisfied with probability one if the probability distribution of noise has no atoms.

Let $\tilde{g} \in C^0(\mathcal{K}, \mathbb{R})$ be a 0-cochain which is a on all the nodes of component 2 and zero on all other nodes. $\delta_0 \tilde{g}$ is either equal to a or $-a$ on all the 1-simplices that connect one component to the other component and is zero on all other 1-simplices. For any a, $g^* - \tilde{g}$ is a feasible solution to the optimization in (5). For $k \in \{1, \ldots, p\}$ let z_k be equal to $(h^* + \delta_0 g^*)([v_i, v_j])$ where $([v_i, v_j])$ is the kth ordered 1-simplex such that v_i is in component 1 and v_j is in component 2. Now consider problem (6) with z_k as defined. Since $h^* + \delta_0 g^*$ is optimal and extreme point for (5), so is $a = 0$ for (6). Hence by Lemma 3, $z_k = 0$ for some $k \in \{1, \ldots, p\}$ which contradicts the fact that the only edges in $\tilde{E}$ are the ones in support of $h^* + \delta_0 g^*$. ∎

If we use an optimization method that always yields an extreme point solution (e.g. the simplex algorithm), we do not need to worry about the graph becoming disconnected by edge deletion.

IV. Problem Reformulation on Consistency Complexes

In the previous section, we orthogonally projected away the locally inconsistent component of the measurements in the first step of the algorithm. This might significantly distort the measurements on the 2-simplices if the local inconsistency is too large. The algorithm we suggest in this section resolves this issue by constructing the two dimensional simplicial complex in a way that the local inconsistencies are absent (or negligible) to start with. In this section instead of working with $F_2(G)$, we introduce a new simplicial complex that we call the *γ-consistent complex* whose topology captures the local consistencies of the measurements.

Definition 6: For any $\gamma \in \mathbb{R}^+ \cup \{+\infty\}$ and pairwise measurements y, the *γ-consistent complex* of graph G, denoted by $\mathcal{K}_\gamma(G, y)$, is an arbitrarily oriented two dimensional simplicial complex whose 1-skeleton is G and which includes $\{v_i, v_j, v_k\}$ as a 2-simplex if and only if $v_i \sim v_j$, $v_j \sim v_k$, $v_k \sim v_i$, and

$$y([v_i, v_j]) + y([v_j, v_k]) + y([v_k, v_i]) \leq \gamma.$$

Remark 2: The definition of γ-consistent complex basically requires 'filling in' a triangle, when the total inconsistency of the measurements on that triangle is no more than γ. This defines a one parameter family of simplicial complexes that allows for calibrating how much local inconsistency we would like to tolerate by changing the value of γ. The γ-consistent complex is a subcomplex of the flag complex. When γ is equal to $+\infty$ we get the two dimensional flag complex used in the previous section. When $\gamma = 0$ we only add the 2-simplices on which the measurements are exactly consistent. Generically this will result in a simplicial complex that has no 2-simplices. Consequently, the first step of our algorithm is trivial for $\gamma = 0$. In such a case our algorithm is the same as (3), the standard estimation problem on a graph, when the norm used is the ℓ^1 norm.

The 'holes' of $\mathcal{K}_\gamma(G, y)$ correspond to regions of G where inconsistencies are large. In a γ-consistent complex the holes are not just artifacts of network structure. Rather, they also capture essential inconsistencies that are caused by excessive noise on certain edges.

All the results of previous section can be readily extended to γ-consistent simplicial complexes by skipping the first step of the algorithm. The resulting estimate is guaranteed to have inconsistency that is no more than γ on any of the 2-simplices.

V. Distributed Implementation

In this section we show that the algorithms proposed in sections III and IV can be implemented by nodes in a distributed way using only local measurements and connectivity information.

In the first two steps we need to orthogonally project y onto $\operatorname{im} \delta_0 \oplus \ker \Delta_1$ and $\ker\ \Delta_1$ respectively. Since we need only the connectivity information to construct δ_0 and Δ_1, this is essentially a local calculation. See [11] for a detailed discussion of how these steps can be carried out in a distributed way.

In the third step we solve the linear program in (5) to get an extreme point solution. Since both the objective and the domain of (5) depend only on graph connectivity, we can use the subgradient [17] or distributed simplex [18] algorithms to solve the program using only local information. Once we get g^*, we can proceed to the fourth step to remove the problematic edges from the graph. This step can easily be implemented in a distributed way.

After deleting the problematic edges, a node which is a neighbor of the anchor node can compute $\hat{x}^*$ by adding the reference node's state to y^* on the edge connecting them. Then it can act as an anchor for its neighbors and so on. Since each connected component has at least one reference node, eventually all the nodes can compute $\hat{x}^*$. All the measurements that are not deleted are (approximately) consistent. Each node gets consistent information from all its neighbors and as a result can compute an estimate $\hat{x}^*$ which is consistent up to γ.

VI. Simulations

We use a random geometric graph as the model of our network (see [9], [19]). x is i.i.d. with a uniform distribution over $[0, 10]$ for all nodes. The measurement noise is i.i.d. Gaussian with variance 1. We solve the optimization problem (5) using simplex algorithm. Some simulation results are presented in Fig. 1. Note that increasing γ corresponds to

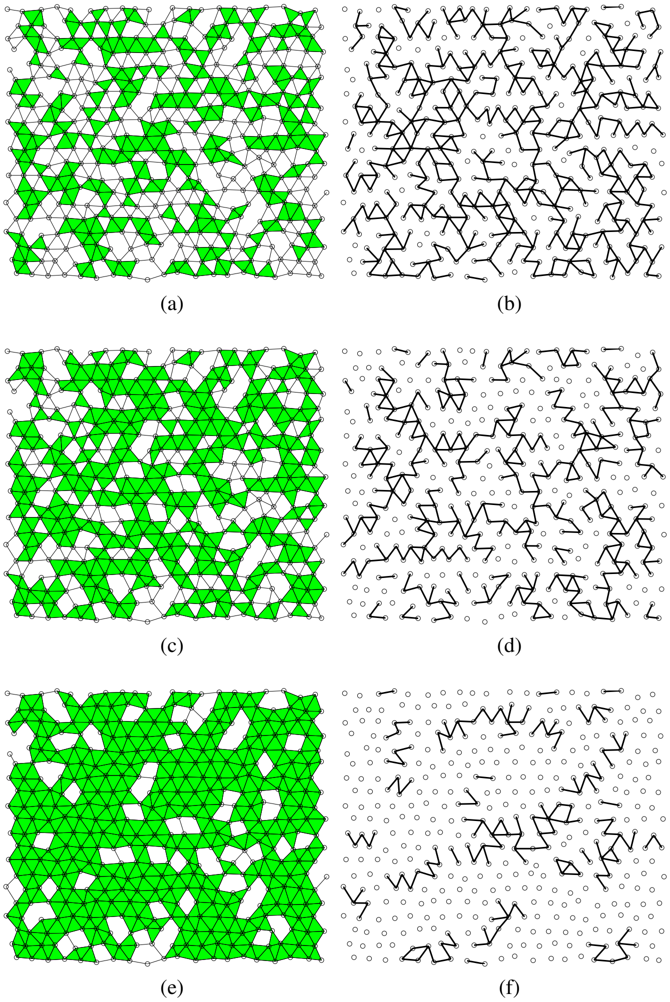

Fig. 1. Simulation results. (a) 0.01-consistent complex (b) deleted edges for $\gamma = 0.01$ (c) 1-consistent complex (d) deleted edges for $\gamma = 1$ (e) $+\infty$-consistent complex (f) deleted edges for $\gamma = +\infty$.

adding the 2-simplices with more inconsistencies to the simplicial complex. Hence, γ-consistent simplicial complexes have fewer 'holes' for higher values of γ. Our algorithm reconstructs x faithfully for a large range of γ.

The simulations clearly demonstrates the tradeoff between the local inconsistencies we are willing to tolerate, and the number of edges we need to discard. When γ is small we get an estimate that is more consistent locally but is obtained by discarding more of the informative measurements, whereas when γ is large the estimate is less consistent locally, but is obtained by using most of the measurements available. This shows that there is an inherent tradeoff between the two main objectives of the problem of estimation from noisy pairwise measurements: small error, that is for the estimate $\hat{y}^*$ to be close to y, and consistency, which requires $\hat{y}^*$ to induce a consistent function on the nodes.

VII. Directions for Future Work

In the future, we would like to explore the tradeoff between the error of an estimate and its consistency more in depth. We would like to see how the total error of our estimate and the number of edges we discard vary when we run our algorithm for different values of γ.

Another possible extension is to use the more general cell complexes instead of simplicial complexes. When cell complexes are used, we can consider the inconsistencies in the more general cycles that have lengths of more than three. For instance if the network is a hexagonal grid, the more natural 2-cells to use would be hexagons. Then we can talk about local inconsistencies on hexagons instead of triangles. Our algorithm will still work in this setting with the caveat that to attach a 2-cell to the cell complex we need the measurements from all the 1-cells that make its boundary. Even though the algorithm is still distributed, but we need information from the nodes with further hop distances.

References

[1] R. Karp, J. Elson, D. Estrin, and S. Shenker, "Optimal and global time synchronization in sensornets," Center for Embedded Networked Sensing, Univ. of California, Los Angeles, Tech. Rep., 2003.

[2] A. Giridhar and P. Kumar, "Distributed clock synchronization over wireless networks: Algorithms and analysis," in *Proc. IEEE CDC'06*, 2006, pp. 4915–4920.

[3] P. Barooah, N. M. D. Silva, and J. P. Hespanha, "Distributed optimal estimation from relative measurements," in *Proc. 3rd Int. Conf. Intelligent Sensing Information Processing (ICISIP), 2005*. Springer-Verlag, 2006, pp. 226–231.

[4] P. Barooah, N. da Silva, and J. Hespanha, "Distributed optimal estimation from relative measurements for localization and time synchronization," in *Distributed Computing in Sensor Systems*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, vol. 4026, pp. 266–281.

[5] W. J. Russell, D. J. Klein, and J. P. Hespanha, "Optimal estimation on the graph cycle space," in *Proc. 2010 American Control Conference (ACC 10)*, Jun 2010.

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

[7] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proce. 4th international symposium on Information processing in sensor networks IPSN'05*, 2005, p. 9.

[8] P. Doyle and E. Snell, "Random walks and electrical networks. Carus Math," 1984.

[9] V. de Silva and R. Ghrist, "Coordinate-free Coverage in Sensor Networks with Controlled Boundaries via Homology," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1205–1222, 2006.

[10] A. Tahbaz-Salehi and A. Jadbabaie, "Distributed coverage verification in sensor networks without location information," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1837–1850, January 2010.

[11] A. Muhammad and M. Egerstedt, "Control using higher order Laplacians in network topologies," in *Proc. 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, 2006.

[12] X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye. (2008) Learning to rank with combinatorial Hodge theory. [Online]. Available: arXiv:0811.1067

[13] A. Hatcher, *Algebraic Topology*. Cambridge University Press, 2002.

[14] J. R. Munkres, *Elements of Algebraic Topology*. Addison-Wesley, 1984.

[15] J. Friedman, "Computing Betti numbers via combinatorial laplacians," *Algorithmica*, vol. 21, pp. 331–346, 1998.

[16] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.

[17] S. Boyd, L. Xiao, and A. Mutapcic. (2003, Oct.) Subgradient methods, lecture notes. [Online]. Available: http://www.stanford.edu/class/ee392o/subgrad_method.pdf

[18] G. Yarmish, "A distributed implementation of the simplex method," Polytechnic University, Brooklin, Tech. Rep., 2001.

[19] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.