

# Optimal Estimation from Relative Measurements: Error Scaling (Extended Abstract)

Prabir Barooah      João P. Hespanha

## I. ESTIMATION FROM RELATIVE MEASUREMENTS

We consider the problem of estimating a number of vector valued variables from a number of noisy “relative measurements”, i.e., measurement of the difference between certain pairs of these variables. This type of measurement model appears in several sensor network problems, such as sensor localization and time synchronization [1]. Consider  $n$  vector-valued variables  $x_1, x_2, \dots, x_n \in \mathbb{R}^k$ , called node variables, one or more of which are known, and the rest are unknown. A number of noisy measurements of the difference  $x_u - x_v$  are available for certain pairs of nodes  $(u, v)$ . We can associate the variables with the nodes  $\mathbf{V} = \{1, 2, \dots, n\}$  of a directed graph  $\mathbf{G} = (\mathbf{V}, \mathbf{E})$  and the measurements with the edges  $\mathbf{E}$  of it, consisting of ordered pairs  $(u, v)$  such that a noisy “relative” measurement between  $x_u$  and  $x_v$  is available:

$$\zeta_{uv} = x_u - x_v + \epsilon_{u,v} \in \mathbb{R}^k, \quad (u, v) \in \mathbf{E} \subset \mathbf{V} \times \mathbf{V}, \quad (1)$$

where the  $\epsilon_{u,v}$ 's are uncorrelated zero-mean noise vectors with known covariance matrices  $P_{u,v} = \mathbb{E}[\epsilon_{u,v}\epsilon_{u,v}^T]$ . Just with relative measurements, determining the  $x_u$ 's is only possible up to an additive constant. To avoid this ambiguity, we assume that there is at least one *reference* node  $o \in \mathbf{V}$  whose node variable  $x_o$  is known. Distributed algorithms to compute the optimal estimate using only local information were reported in [1], where the *Optimal estimate* refers to the classical Best Linear Unbiased (BLU) Estimator, which achieves the minimum variance among all linear unbiased estimators.

## II. THE QUESTION OF ERROR SCALING

One may wonder what are the *fundamental limitations of estimation accuracy* for truly large graphs. Reasons for concern arise from estimation problems such as the one associated with the simple graph shown in Figure 1. It is a chain of nodes with node 1 as the reference and with a single edge  $(u+1, u)$ ,

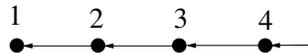


Fig. 1. A graph where  $x_u$ 's optimal estimate has a error variance that grows linearly with distance from the reference.

$u \geq 1$  between consecutive nodes  $u$  and  $u+1$ . Without much difficulty, one can show that for such a graph the optimal estimate of  $x_u$  is given by

$$\hat{x}_u = \zeta_{u,u-1} + \dots + \zeta_{3,2} + \zeta_{2,1} + x_1,$$

and since each measurement introduces an additive error, the variance of the optimal estimation error  $\hat{x}_u - x_u$  will increase linearly with  $u$ . This means that if  $u$  is very “far” from the reference node 1 then its estimate will necessarily be quite poor. Although the precise estimation error depends on the

This material is based upon work supported by the Institute for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office and by the National Science Foundation under Grant No. CCR-0311084.

Both authors are with the Dept. of Electrical and Computer Engineering and the Center for Control, Dynamical-Systems, and Computation at Univ. of California, Santa Barbara, CA 93106.

exact values of the variances of the measurements, for this graph the variance of the optimal estimate of  $x_u$  will grow essentially linearly with  $u$ .

We investigate how the structure of the graph  $\mathbf{G}$  affects the “quality” of the optimal estimate  $\hat{x}_u^*$  of  $x_u$ , measured in terms of the covariance of the estimation error  $\Sigma_{u,o} := \mathbb{E}[(x_u - \hat{x}_u^*)(x_u - \hat{x}_u^*)^T]$ . We focus on the case of a single reference node. Specifically, we raise and answer the following questions:

- 1) How does the error variance of a node variable’s estimate scale with the node’s distance from the reference node asymptotically, and
- 2) can one deduce these scaling laws from coarse structure of the measurement graph  $\mathbf{G}$ ?

For a given maximum acceptable error, the number of nodes with acceptable estimation errors will be large if the graph exhibits a slow increase of variance with distance, but small otherwise. These scaling laws therefore determine how large a graph can be in practice. We will describe here a classification of graphs that determines how the variance grows with distance from the reference node. There are graphs where variance grows linearly with distance, as in the example of Figure 1. But there are also a large class of graphs where it grows only logarithmically with distance. Most surprisingly, in certain graphs it can even stay below a constant value, no matter the distance. Our results also point out inadequacies in conventional measures of graph denseness, such as node degree, in predicting how estimation accuracy scales with distance.

### III. ELECTRICAL ANALOGY

It can be shown that the error covariance matrices of the BLU estimator are numerically equal to the effective resistances in an appropriately defined *generalized* resistive electrical network where currents, potentials and resistances are matrices. Resistances in such a network are always square positive definite matrices in  $\mathbb{R}^{k \times k}$  and are called *generalized resistances*. Currents and potentials are matrices in  $\mathbb{R}^{k \times k'}$ , where  $k' \leq k$ , and are called *generalized currents* and *generalized potentials*. Given a measurement graph  $\mathbf{G}$  with measurement error covariances  $P : \mathbf{E} \rightarrow \mathbb{R}^{k \times k}$ , we form an electrical network with the graph  $\mathbf{G}$  and generalized resistances  $R(e) = P(e)$  for every edge  $e \in \mathbf{E}$ .

The following result proved in [2] establishes the connection between optimal estimator covariance and generalized effective resistance.

**Theorem 1 (Electrical Analogy).** *Consider a measurement graph  $\mathbf{G}$  with a single reference node  $o$  and construct a generalized electric network with  $k \times k$  edge-resistors that are numerically equal to the covariance matrices of the edge-measurement errors. For every node  $u$ , the  $k \times k$  covariance matrix  $\Sigma_u$  of the estimation error of  $x_u$  is equal to the generalized effective resistance between the node  $u$  and the reference node  $o$ .  $\square$*

When  $k = 1$ , i.e., variables and measurements are scalar-valued, the generalized electrical network is the familiar, “regular” electrical network, and the generalized effective resistance is the scalar-valued effective resistance. In [4], it was shown that the variance of the optimal estimate of time shifts between two clocks is the effective resistance between the respective nodes in the measurement graph. The theorem above is a generalization of this result to the case of vector-valued variables.

The matrix-valued effective resistance in a graph can be bounded by that in another graph if one graph can be *embedded* in another graph. Consider two graphs  $\mathbf{G} = (\mathbf{V}, \mathbf{E})$  and  $\bar{\mathbf{G}} = (\bar{\mathbf{V}}, \bar{\mathbf{E}})$  such that  $\mathbf{V} \in \bar{\mathbf{V}}$  and whenever two nodes  $u$  and  $v$  in  $\mathbf{V}$  have an edge between them in  $\mathbf{E}$ , they also have an edge in  $\bar{\mathbf{E}}$ . In that case, we say that  $\mathbf{G}$  is embedded in  $\bar{\mathbf{G}}$  or that  $\bar{\mathbf{G}}$  embeds  $\mathbf{G}$ . The following theorem proved in [2] shows how the effective resistances in two graphs, when one is embedded in another, are related. Given two symmetric matrices  $A$  and  $B$ , we write  $A \geq B$  to mean that  $A - B$  is a positive semi-definite matrix.

**Theorem 2 (Rayleigh’s Generalized Monotonicity Law).** *Consider two generalized electrical networks with graphs  $\mathbf{G}$  and  $\bar{\mathbf{G}}$  such that  $\mathbf{G}$  can be embedded in  $\bar{\mathbf{G}}$ , and the matrix edge resistances in*

$\mathbf{G}$  are greater than the corresponding edge resistances in  $\bar{\mathbf{G}}$ . Then for every pair of nodes  $u, v \in \mathbf{V}$  of  $\mathbf{G}$ , we have  $R_{u,v}^{\text{eff}} \geq \bar{R}_{\bar{u},\bar{v}}^{\text{eff}}$ , where  $R_{u,v}^{\text{eff}}$  denotes the effective resistance between  $u$  and  $v$ , and  $\bar{R}_{\bar{u},\bar{v}}^{\text{eff}}$  denotes the effective resistance between the corresponding nodes  $\bar{u}$  and  $\bar{v}$  in  $\bar{\mathbf{G}}$ .  $\square$

The name of the theorem comes from a result known for “regular” electrical networks as *Rayleigh’s Monotonicity Law*, which states that the effective resistance between any two nodes of an electrical network is a monotonic function of the edge resistances [3]. The theorem above can be viewed as a generalization to the case of matrix-valued effective resistance.

#### A. Lattices, $h$ -fuzzes, and their effective resistance

A  $d$ -dimensional lattice, denoted by  $\mathbf{Z}_d$  is a graph that has a vertex at every point in  $\mathbb{R}^d$  with integer coordinates and an edge between every two vertices with an Euclidean distance between them equal to one. Edge directions are arbitrary. Effective resistance in lattices, and in a class of graphs derived from them called lattice fuzzes, are especially useful in studying the scaling laws of effective resistance, and therefore covariance, in large graphs.

Given a graph  $\mathbf{G}$  and an integer  $h \geq 1$ , the  $h$ -fuzz of  $\mathbf{G}$ , denoted by  $\mathbf{G}^{(h)}$ , is a graph with the same set of nodes as  $\mathbf{G}$  but with a larger set of edges. In particular,  $\mathbf{G}^{(h)}$  has an edge between  $u$  and  $v$  whenever the graphical distance between  $u$  and  $v$  is less than or equal to  $h$  [3]. The directions of the “new” edges are arbitrary (see the comment following Theorem 1). The following lemma establishes the effective resistance of a  $d$ -D lattice and its  $h$ -fuzz. For a proof of this result, see [2].

**Lemma 1 (Lattice Effective Resistance).** *Consider the electrical network constructed from the  $h$ -fuzz of the  $d$ -D lattice by assigning a constant generalized resistance  $R_o \in \mathbb{R}^{k \times k}$  on every edge of  $\mathbf{Z}_d^{(h)}$ , where  $h$  is a positive integer. We denote by  $d_{\mathbf{Z}_d}(u, v)$  the graphical distance between two nodes  $u$  and  $v$  in the  $d$ -D lattice  $\mathbf{Z}_d$ . The effective resistance  $R_{u,v}^{\text{eff}}$  between two nodes  $u$  and  $v$  in the electrical network  $\mathbf{Z}_d^{(h)}$  satisfies (1)  $R_{u,v}^{\text{eff}}(\mathbf{Z}_1^{(h)}) = \Theta(d_{\mathbf{Z}_1}(u, v))$ , (2)  $R_{u,v}^{\text{eff}}(\mathbf{Z}_2^{(h)}) = \Theta(\log d_{\mathbf{Z}_2}(u, v))$ , and (3)  $R_{u,v}^{\text{eff}}(\mathbf{Z}_3^{(h)}) = \Theta(1)$ .  $\square$*

The usual asymptotic notation  $\Theta(\cdot)$  is used with matrix valued functions in the following way. For two functions  $g : \mathbb{R} \rightarrow \mathbb{R}^{k \times k}$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$ , the notation  $g(x) = \Theta(f(x))$  means there exists a constant  $x_o$  and two positive definite matrices  $A$  and  $B$  such that  $Af(x) \leq g(x) \leq Bf(x)$  for all  $x > x_o$ . The notations  $\Omega(\cdot)$  and  $\mathcal{O}(\cdot)$  will be similarly used with matrix valued functions later on.

## IV. GRAPH DRAWING

In graph theory, a graph is generally treated purely as a collection of nodes connected by edges, without any regard to the geometry determined by the nodes’ locations. However, for the graphs that arise in sensor networks there is an underlying geometry because nodes generally correspond to physical agents and their locations often determine the connectivity of the graph. Graph drawings are used to capture the geometry of graphs in Euclidean space.

The drawing of a graph  $\mathbf{G} = (\mathbf{V}, \mathbf{E})$  is simply a mapping of its nodes to points in some Euclidean space, which can formally be described by a function  $f : \mathbf{V} \rightarrow \mathbb{R}^d$ ,  $d \geq 1$ . Given two nodes  $u, v \in \mathbf{V}$  the *Euclidean distance between  $u$  and  $v$  induced by the drawing  $f : \mathbf{V} \rightarrow \mathbb{R}^d$*  is defined by  $d_f(u, v) := \|f(v) - f(u)\|$ , where  $\|\cdot\|$  denoted the usual Euclidean norm in  $d$ -space. Euclidean distances depend on the drawing and can be completely different from graphical distances. It is important to emphasize that the definition of drawing does not require edges to not intersect and every graph has a drawing in any Euclidean space.

For graphs that arise in the estimation problems in sensor networks, there is a *natural drawing* that is obtained by associating each node to its position in 1-, 2- or 3-dimensional Euclidean space. In reality, all sensors are situated in 3-dimensional space. However, sometimes it maybe more natural to

draw them on a 2-dimensional Euclidean space if one dimension (e.g., height) does not vary much from node to node, or is somehow irrelevant. For *natural drawings*, the Euclidean distance induced by the drawing is, in general, a much more meaningful notion of distance than the graphical distance. The Euclidean distance induced by an appropriate drawings provide the right measure of distance to determine scaling laws of effective resistance.

1) *Dense graphs*: We define a graph to be *dense in  $\mathbb{R}^d$*  if its nodes can be drawn in  $\mathbb{R}^d$  such that there is a positive number  $\gamma$  so that (i) every ball in  $\mathbb{R}^d$  with diameter  $\gamma$  contains at least one node of the graph and (ii) there is a nonzero minimum ratio  $\rho$  of the Euclidean distance in the drawing to the graphical distance, between any two nodes of the graph. Intuitively, dense graphs in  $\mathbb{R}^d$  have sufficiently many nodes to cover  $\mathbb{R}^d$  without holes, and sufficiently many interconnecting edges so that two nodes with a small Euclidean distance in the drawing also has a small graphical distance. Intuitively, these graphs are “dense” in the sense that the nodes can cover  $\mathbb{R}^d$  without leaving large holes between them and still having sufficiently many edges so that a small Euclidean distance between two nodes in the drawing guarantees a small graphical distance between them. In particular, for dense drawings there are always finite constants  $\alpha, \beta$  such that [2]

$$d_{\mathbf{G}}(u, v) \leq \alpha d_f(u, v) + \beta, \quad \forall u, v \in \mathbf{V}.$$

2) *Sparse graphs*: We define a graph to be *sparse in  $\mathbb{R}^d$*  (for some  $d$ ) if it can be drawn in  $\mathbb{R}^d$  in a *civilized* manner. Graphs that can be drawn in a civilized manner appeared in [3] in connection with random walks, where it was defined as a graph that can be drawn in  $\mathbb{R}^d$  such that (i) there is a minimum distance  $s > 0$  between any two nodes and (ii) there is a maximum distance  $r < \infty$  between nodes connected by an edge in the drawing. Intuitively, the nodes and edges of such a graph are sufficiently sparse to be drawn in  $\mathbb{R}^d$  without too much clutter.

In sensor networks natural drawings generally provide a good starting point to determine whether a graph is sparse or dense in some Euclidean space. For example, we can conclude from the natural drawing of a  $d$ -dimensional lattice that this graph is both dense and sparse in  $\mathbb{R}^d$ . One can also show that a  $d$ -dimensional lattice can never be dense in  $\mathbb{R}^{\bar{d}}$  with  $\bar{d} > d$ . This means, for example, that any drawing of a 2-dimensional lattice in the 3-dimensional Euclidean space will never be dense. Moreover, a  $d$ -dimensional lattice can never be drawn in a civilized way in  $\mathbb{R}^{\bar{d}}$  with  $\bar{d} < d$ . This means, for example, that any drawing of a 3-dimensional lattice in the 2-dimensional Euclidean space will never be a civilized drawing. A 3-dimensional lattice is therefore not sparse in  $\mathbb{R}^2$ .

The notions of graph “sparseness” and “denseness” are mostly interesting for infinite graph, because every finite graph is sparse in all Euclidean spaces  $\mathbb{R}^d, \forall d \geq 1$  and no finite graph can ever be dense in any Euclidean space  $\mathbb{R}^d, \forall d \geq 1$ . However, in practice infinite graphs serve as proxies for very large graphs that, from the perspective of most nodes, “appear to extend in all directions as far as the eye can see.” So conclusions drawn for sparse/dense infinite graphs hold for large graphs, at least far from the graph boundaries.

The notions of sparseness and denseness introduced above are useful because they provide a complete characterization for the classes of graphs that can embed or be embedded in lattices, for which the Lattice Effective Resistance Lemma 1 provides the precise scaling laws for the effective resistance.

**Theorem 3 (Dense/Sparse Embedding).** *Let  $\mathbf{G} = (\mathbf{V}, \mathbf{E})$  be a graph without multiple edges between the same pair of nodes.*

- 1)  $\mathbf{G}$  is sparse in  $\mathbb{R}^d$  if and only if  $\mathbf{G}$  can be embedded in an  $h$ -fuzz of a  $d$ -dimensional lattice.
- 2)  $\mathbf{G}$  is dense in  $\mathbb{R}^d$  if and only if (i) the  $d$ -dimensional lattice can be embedded in an  $h$ -fuzz of  $\mathbf{G}$  for some positive integer  $h$  and (ii) every node of  $\mathbf{G}$  is at an uniformly bounded graphical distance from another node of  $\mathbf{G}$  that is also a node of  $\mathbf{Z}_d$ .  $\square$

TABLE I

EFFECTIVE RESISTANCES FOR GRAPHS THAT ARE SPARSE OR DENSE. THE NOTATIONS  $\Omega(\cdot)$  AND  $\mathcal{O}(\cdot)$  ARE USED WITH MATRIX VALUED QUANTITIES AS DESCRIBED AFTER LEMMA 1.

Euclidean space	Covariance matrix of the estimation error of $x_u$ in a <i>sparse graph</i>	Covariance matrix of the estimation error of $x_u$ in a <i>dense graph</i>
 $\mathbb{R}$	$R_{u,o}^{\text{eff}} = \Omega(d_f(u, o))$	$R_{u,o}^{\text{eff}} = \mathcal{O}(d_f(u, o))$
 $\mathbb{R}^2$	$R_{u,o}^{\text{eff}} = \Omega(\log d_f(u, o))$	$R_{u,o}^{\text{eff}} = \mathcal{O}(\log d_f(u, o))$
 $\mathbb{R}^3$	$R_{u,o}^{\text{eff}} = \Omega(1)$	$R_{u,o}^{\text{eff}} = \mathcal{O}(1)$

The first statement of the lemma is essentially taken from [3]; the rest are proved in [2]. The assumption about there being no parallel edges is not restrictive since parallel edges can be replaced by a single edge with an equivalent resistance without changing the effective resistances.

## V. SCALING LAWS FOR VARIANCE

We are now finally ready to characterize scaling laws for the optimal Estimator covariance in terms of the denseness/sparseness properties of the graph. The following theorem from [2] does precisely this by combining the Electrical Analogy Theorem 1, Rayleigh’s Generalized Monotonicity Law, the Lattice Effective Resistance Lemma 1, and the Dense/Sparse Embedding Theorem 3.

**Theorem 4 (Scaling of effective resistance).** *Consider a measurement graph  $\mathbf{G} = (\mathbf{V}, \mathbf{E})$  with measurement error covariances  $P : \mathbf{E} \rightarrow \mathbb{R}^{k \times k}$  that satisfy  $P_{\min} \leq P(e) \leq P_{\max}$ ,  $\forall e \in \mathbf{E}$  for some symmetric positive definite matrices  $P_{\min}$ ,  $P_{\max}$ . If the graph  $\mathbf{G}$  is either sparse or dense in some Euclidean space  $\mathbb{R}^d$ , then the formulas in Table I gives the bounds on the asymptotic scaling of optimal estimator error covariance  $\Sigma_{u,o}$  for every node  $u \in \mathbf{V}$  with  $o \in \mathbf{V}$  as the reference. In the table,  $d_f(u, o)$  denotes the Euclidean distance between node  $u$  and the reference node  $o$ , for any drawing  $f$  that establishes the graph’s sparseness/denseness.  $\square$*

### A. Are Sensor Networks Dense/Sparse:

One may ask whether it is common for the graphs that arise in distributed control/estimation problems to be sparse and/or dense in some Euclidean space  $\mathbb{R}^d$ . The answer happens to be “very much so” and this is often seen by considering the natural drawing of the graph. Recall that a natural drawing associates each node with its physical position in 1-, 2-, or 3-dimensional Euclidean space (cf. discussion in Section IV). All natural drawings are likely to be sparse in 3-dimensional space, since the only requirements for sparseness are that nodes not lie on top of each other and edges be of finite length. When nodes in a 2-dimensional domain or when the third physical dimension is irrelevant, again the natural drawing is likely to be civilized in 2-dimensional space for the same reasons. It is slightly harder for a graph to satisfy the denseness requirements. Formally, a graph has to be infinite to be dense. However, what matters in practice are the properties of the graph “not too close to the boundary”. Thus, a large graph satisfies the denseness requirements, as long as there are no big holes between nodes and sufficiently many interconnections between them. For example, the commonly encountered model consisting of nodes that are Poisson distributed random points in 2-dimensional space with an edge between every pair of nodes that they are within a certain range are likely to be dense in 2-dimensional for a sufficiently large range (when compared to the intensity of the Poisson process). In any case, almost all graphs that appear in distributed control/estimation problems are likely to fall into at least one of the classes - sparse or dense in some  $\mathbb{R}^d$ ,  $1 \leq d \leq 3$ .

## B. Counterexamples to Conventional Wisdom

An interesting fallout of our results is the way they point out the inadequacies in conventional measures of graph “denseness”, such as node degree or density of nodes/edges per unit area for predicting how estimation error scales with distance. As an example, Figure 2 shows two graphs – a triangular lattice and a 3-fuzz of a 1-dimensional lattice. Applying the results in this paper, we can readily establish that the effective resistance in the 3-fuzz of the 1-dimensional lattice grows linearly with distance, whereas in the triangular lattice it grows only with the logarithm of distance, in spite of both graphs having the *same node degree of 6*. Other interesting counterexamples to conventional wisdom are described in [2].

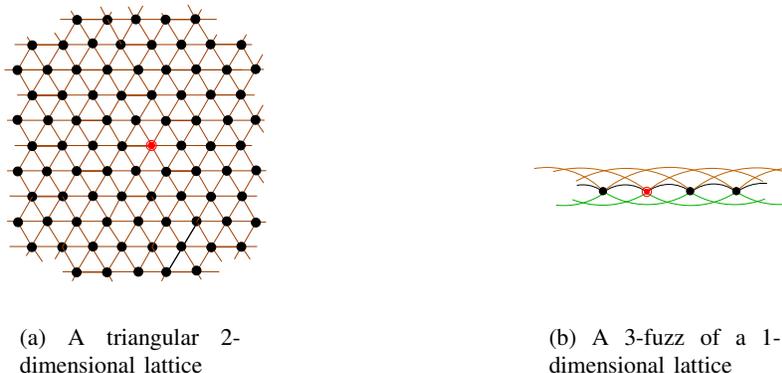


Fig. 2. Two different measurement graphs — a triangular 2-dimensional lattice and a 3-fuzz of a 1-dimensional lattice. Both graphs have the same node degree for every node but very different variance growth rates with distance.

## VI. CONCLUSION

We have established a classification of graphs – dense and sparse – that determine how the optimal estimator variance scales with distance. If a graph is dense in some Euclidean space, we can establish upper bounds on the variance of a node variable’s estimate as a function of the node’s distance from the reference node in an appropriate drawing. For sparse graphs we get lower bounds. These scaling laws, being true for the optimal estimate, determine fundamental limitations on the accuracy achievable by any estimation algorithm. Moreover, they show what structural properties of a graph determine the scaling of the variance, and can help us design networks where more accurate estimates are possible. Our results also point out inadequacies of traditional measures of graph density in predicting scaling of estimation errors.

## REFERENCES

- [1] P. Barooah, N. M. da Silva, and J. P. Hespanha. Distributed optimal estimation from relative measurements in sensor networks: Applications to localization and time synchronization. Accepted for publication in DCOSS’06, June 2006.
- [2] P. Barooah and J. P. Hespanha. Optimal estimation from relative measurements: Electrical analogy and error bounds. Technical report, University of California, Santa Barbara, 2003. URL <http://www.ccec.ece.ucsb.edu/~pbarooah/publications/TR1.html>.
- [3] P. G. Doyle and J. L. Snell. Random walks and electric networks. Math. Assoc. of America, 1984.
- [4] R. Karp, J. Elson, D. Estrin, and S. Shenker. Optimal and global time synchronization in sensornets. Technical report, Center for Embedded Networked Sensing, Univ. of California, Los Angeles, 2003.