DISTRIBUTED GRADIENT ESTIMATION USING RANDOM SENSOR NETWORKS

by

Slobodan N. Simić and Shankar Sastry

Memorandum No. UCB/ERL M02/25

6 September 2002

COVER

DISTRIBUTED GRADIENT ESTIMATION USING RANDOM SENSOR NETWORKS

by

Slobodan N. Simić and Shankar Sastry

Memorandum No. UCB/ERL M02/25

6 September 2002

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

DISTRIBUTED GRADIENT ESTIMATION USING RANDOM SENSOR NETWORKS

SLOBODAN N. SIMIĆ AND SHANKAR SASTRY

ABSTRACT. We present a distributed algorithm for estimating the gradient of an environmental scalar field (such as temperature, the intensity of light, atmospheric pressure, etc.) using a random sensor network. We derive an error estimate, discuss the algorithm's complexity, and show some simulation results. Potential applications are in preventing forest fires, energy conservation, oceanography, building science, etc.

INTRODUCTION

Recent advances in MEMS, computing, and communication technology have sparked the emergence of massively distributed, wireless sensor networks consisting of hundreds or potentially thousands of nodes. Each node is able to sense the environment, perform simple computations, and communicate with its peers or to an external observer. The challenges these networks present are beyond the reach of current theory and algorithms.

In this paper, we present a distributed algorithm for monitoring an environmental scalar field (such as temperature, intensity of light, atmospheric pressure, etc.) using a random wireless sensor network. We also derive an error estimate in terms of the parameters of the sensor network. Possible applications of the algorithm are in preventing forest fires, energy conservation, oceanography, building science, etc.

Our work is mainly motivated by the Sensorwebs and Smart Dust [KKP] projects at UC Berkeley, whose aim is to develop a unified framework for distributed sensor networks. Some previous work on environmental monitoring using random sensor networks was done in [Doh00].

Due to high long range communication costs and low battery power, it is natural to seek decentralized, distributed algorithms for sensor networks. This means that instead of relaying data to a central location which does all the computing, the nodes process information in a collaborative, distributed way. For instance, they can form computational clusters, based on their distance from each other. The outcome of these distributed, local computations is stored in local memory and can then be, when necessary, relayed to a centralized computing

Date: September 5, 2002.

Partially supported by the NSF grant ITR/SI: Societal Scale Information Systems: Technologies, Design and Applications, number EIA-0122599.

unit. Robustness to node failures is another reason to seek distributed rather than centralized algorithms.

The basic idea of our algorithm is the following. Each node communicates with its neighbors and computes the maximal difference quotient of the measured variable. The estimate of the gradient at each node is taken to be the vector in the corresponding direction with norm equal to the maximal difference quotient. The algorithm is simple enough to be implemented on the current sensor network platform of Mica motes [Cul].

We point out that the main purpose of this paper is to rigorously analyze the accuracy and complexity of our algorithm from a probabilistic point of view; its purpose is *not* to deal with technical details of time synchronization, data fusion, communication protocols, etc. We believe that this approach fills a void in the literature as most articles known to the authors do not rigorously address mathematical aspects of sensor networks.

The paper is organized as follows. In Section 1, we introduce the terminology, notation, and the environmental monitoring problem. Section 2 describes the algorithm. In Section 3, we derive an error estimate; Section 4 discusses average complexity, followed by some simulation results in Section 5. The paper concludes with a summary of the results and discussion of future work.

1. Preliminaries

In this section we introduce the basic mathematical framework and formulate the problem.

Assume that a random sensor network consisting of N nodes $S_1, \dots, S_N$ is deployed in some region $D \subset \mathbb{R}^2$. The number i will be called the ID of the node S_i. Assume that:

- Every node is aware of its own position p_i in some fixed coordinate system in D. That is, the network is assumed to have performed node localization (see, e.g., our earlier work [SS01]).
- Each node S_i measures some environmental scalar field V such as temperature, pressure, or the amount of light at its own location. We assume that its measurement v_i is exact, i.e., $v_i = V(p_i)$.
- Each node has a maximal isotropic RF communication range R, i.e., two nodes can communicate if they are less then R meters apart. For every $0 < r \leq R$, each node can adjust it signal strength to achieve communication range r.

Remark. If $0 < r \leq R$ is the maximal communication range, then any two nodes whose distance is $\leq r$ are called *r-neighbors*.

The *goal* is:

> Using only the information collected by the sensor network, design a distributed algorithm for estimating the gradient of V at $p_1, \dots, p_N$.

We make the following assumptions on D, V, and the network.

- D has unit area and is homeomorphic to the closed unit disk in $\mathbb{R}^2$;
- $V : W \to \mathbb{R}$ is a function of class C^2, i.e., twice continuously differentiable, where W is some neighborhood of D in $\mathbb{R}^2$.
- For each $1 \leq i \leq N$, p_i is a random variable with uniform distribution on D.

Notation. Throughout this paper, $\cdot$ will denote the standard dot product on $\mathbb{R}^2$. The corresponding 2-norm of a vector $v \in \mathbb{R}^2$ is $|v| = \sqrt{v \cdot v}$. For a matrix $A \in \mathbb{R}^{2\times 2}$, $\|A\|$ will denote its operator norm relative to $|\ |$,

$$\|A\| = \sup\{|Av| : v \in \mathbb{R}^2, |v| = 1\}.$$

Further, for $a, b \in D$, $a \neq b$, denote the difference quotient of V at a relative to b by

$$Q(a,b) = \frac{V(b) - V(a)}{|b-a|}.$$

Finally, let

$$G(a,b) = Q(a,b)\frac{b-a}{|b-a|}.$$

2. The algorithm

Let $S = S_i$ for some $1 \leq i \leq N$ be a node with position $p = p_i$. Assume the signal strength of all the nodes has been adjusted to achieve maximum communication range of r meters.

We now state our algorithm for estimating $\nabla V(p)$, called $GRAD_S(r)$.

Step 1: *INITIALIZE variables:* $q(S) = 0, n(S) = i$.

Step 2: *SEND "Hello, what is your ID, position, and measurement?"*

Each r-neighbor S_ν replies with $M_\nu = (\nu, p_\nu, v_\nu)$, where ν is its ID, p_ν its position, and v_ν its measurement of V at p_ν.

Step 3: *For each r-neighbor ν, COMPUTE $Q(p, p_\nu)$.*

If $Q(p, p_\nu) > q(S)$ then

$n(S) = \nu$, $q(S) = Q(p, p_\nu)$.

Step 4: *STOP when all responses have been processed. The estimate of $\nabla V(p)$ is*

$$\mathrm{Grad}(p) = G(p, p_{n(S)}).$$

Note that $v_\nu = V(p_\nu)$.

Remark. The algorithm maximizes the difference quotient $Q(p, p_\nu)$ over all neighbors S_ν of S. $\mathrm{Grad}(p)$ is the vector parallel to $p_{n(S)} - p$ of length $Q(p, p_{n(S)})$.

Observe that the algorithm is distributed. The number of operations it executes is a constant multiple of the number of r-neighbors of S. The only operations a node needs to be able to perform are the four elementary arithmetic operations, squaring, square root, and comparisons.

3. Error estimates for $GRAD_S(r)$

We will need the following estimate. Here $\angle(u,v)$ will denote the angle between vectors $u, v \in \mathbb{R}^2$.

3.1. **Proposition.** *For all $p, q \in D$, $p \neq q$,*

$$|G(p,q) - \nabla V(p)| \leq |\nabla V(p)| \sin|\angle(\nabla V(p), q-p)| + \frac{1}{2}\|D^2V\|_\infty |q-p|.$$

Proof. By the Fundamental Theorem of Calculus,

$$V(q) - V(p) = \nabla V(p)\cdot(q-p) + \frac{1}{2}D^2V(\xi)(q-p)\cdot(q-p),$$

for some ξ lying on the segment connecting p and q. Therefore,

$$\begin{aligned} |G(p,q) - \nabla V(p)| &= \left|\frac{V(q)-V(p)}{|q-p|^2}(q-p) - \nabla V(p)\right| \\ &\leq \left|\frac{\nabla V(p)\cdot(q-p)}{|q-p|^2}(q-p) - \nabla V(p)\right| + \frac{1}{2}\left|\frac{D^2V(\xi)(q-p)\cdot(q-p)}{|q-p|^2}(q-p)\right| \\ &= \mathrm{I} + \mathrm{II}. \end{aligned}$$

Consider first

$$\mathrm{I} = \frac{|[\nabla V(p)\cdot(q-p)](q-p) - |q-p|^2\nabla V(p)|}{|q-p|^2}. \tag{1}$$

Letting $v = \nabla V(p)$ and $x = q-p$, by elementary linear algebra we obtain that the numerator of (1) is

$$\begin{aligned} |(v\cdot x)x - |x|^2 v| &= \{[(v\cdot x)x - |x|^2 v]\cdot[(v\cdot x)x - |x|^2 v]\}^{1/2} \\ &= |x|^2|v|\sin|\angle(v,x)| \end{aligned}$$

Thus,

$$\mathrm{I} = |\nabla V(p)|\sin|\angle(\nabla V(p), q-p)|.$$

It is not hard to see that

$$\mathrm{II} \leq \frac{1}{2}\|D^2V\|_\infty |q-p|.$$

This completes the proof of the Proposition. □

For every $1 \leq i \leq N$, denote by θ_i the angle between $\nabla V(p_i)$ and the vector $p_{n(S_i)} - p_i$ (Fig. 1).

3.2. **Corollary.** *For every $1 \leq i \leq N$,*

$$|\mathrm{Grad}(p_i) - \nabla V(p_i)| \leq |\nabla V(p_i)|\sin|\theta_i| + \frac{1}{2}\|D^2V\|_\infty |p_{n(S_i)} - p_i|.$$

Let

$$A_i = |\nabla V(p_i)| \quad \text{and} \quad H = \|D^2V\|_\infty.$$

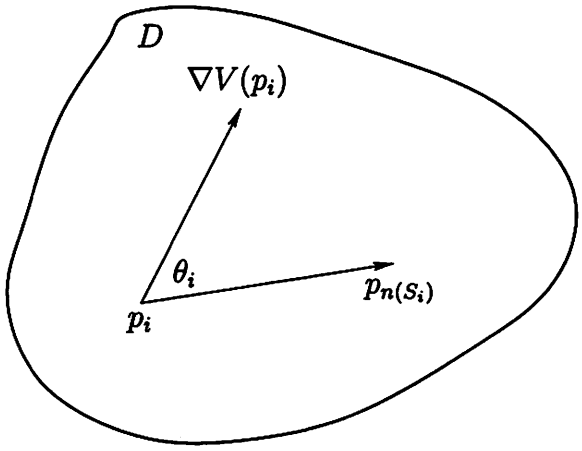

FIGURE 1. The angle θ_i.

3.3. **Lemma.** *Let $q, q_1, \ldots, q_K \in D$ be distinct points and let*

$$\alpha_i = |\angle(\nabla V(q), q_i - q)|.$$

There exist $\rho > 0$ such that for all q_i, q_j with $|q_i - q|, |q_j - q| < \rho$ and $\alpha_i, \alpha_j < \pi/2$, the following holds:

$$\alpha_i < \alpha_j \Rightarrow Q(q, q_i) > Q(q, q_j).$$

In other words, in a sufficiently small polar coordinate neighborhood of q, $q_i \mapsto Q(q, q_i)$ decreases as α_i increases.

Proof. Let

$$c = \min\{|\cos\alpha_m - \cos\alpha_n| : \alpha_m \neq \alpha_n,\ \alpha_m, \alpha_n < \pi/2,\ 1 \leq m, n \leq K\}.$$

Since $c > 0$, we can choose $\rho > 0$ so that

$$\rho < \min\left\{\frac{A_i c}{2H} : 1 \leq i \leq K\right\}.$$

Assume $|q_i - q|, |q_j - q| < \rho$, $\alpha_i, \alpha_j < \pi/2$, and $\alpha_i < \alpha_j$. Then

$$\begin{aligned}
Q(q, q_i) - Q(q, q_j) &= \left\{\nabla V(q)\frac{q_i - q}{|q_i - q|} + \frac{1}{2}D^2V(\xi_i)(q_i - q)\cdot\frac{q_i - q}{|q_i - q|}\right\} \\
&\quad - \left\{\nabla V(q)\frac{q_j - q}{|q_j - q|} + \frac{1}{2}D^2V(\xi_j)(q_j - q)\cdot\frac{q_j - q}{|q_j - q|}\right\} \\
&= |\nabla V(q)|(\cos\alpha_i - \cos\alpha_j) \\
&\quad + \frac{1}{2}\left\{D^2V(\xi_i)(q_i - q)\cdot\frac{q_i - q}{|q_i - q|} - D^2V(\xi_j)(q_j - q)\cdot\frac{q_j - q}{|q_j - q|}\right\}, \\
&= \mathrm{I} + \mathrm{II},
\end{aligned}$$

where ξ_i is a point on the segment connecting q and q_i, and similarly for ξ_j. Further, $|\mathrm{II}| \le 2\rho H$, and $\mathrm{I} \ge A_i c$. Therefore,

$$\mathrm{I} + \mathrm{II} \ge \mathrm{I} - |\mathrm{II}| \ge A_i c - 2\rho H > 0,$$

implying $Q(q, q_i) > Q(q, q_j)$. □

Denote by $\mathrm{P}(A|B)$ and $E(A|B)$ the conditional probability and expectation of A given B [GS97]. Let ∂D be the boundary of D, and $d(x, \partial D)$ the distance from x to ∂D.

3.4. **Proposition.** *For all $1 \le i \le N$ and $\epsilon > 0$ small enough,*

$$\mathrm{P}(|\mathrm{Grad}(p_i) - \nabla V(p_i)| < \epsilon \mid d(p_i, \partial D) \ge r) \ge 1 - [1 - \mu_V(\epsilon)]^{N-1},$$

where

$$\mu_V(\epsilon) = \max\{u_2^2 \sin^{-1} u_1 : A_i u_1 + H u_2 = \epsilon,\ u_1, u_2 > 0\}.$$

In particular, if p_i is an equilibrium of ∇V, then

$$\mathrm{P}(|\mathrm{Grad}(p_i)| < \epsilon \mid d(p_i, \partial D) \ge r) \ge 1 - \left(1 - \frac{4\pi\epsilon^2}{H^2}\right)^{N-1}.$$

Proof. Let $C_i(u_1, u_2)$ be the circular sector at p_i of radius $u_2 > 0$, angular width $\sin^{-1} u_1$ ($u_1 > 0$), and axis of symmetry $\nabla V(p_i)$. If $A_i u_1 + H u_2 < \epsilon$, then by Corollary 3.2,

$$\mathrm{P}(|\mathrm{Grad}(p_i) - \nabla V(p_i)| < \epsilon \mid d(p_i, \partial D) \ge r) \ge \mathrm{P}(p_{n(S_i)} \in C_i(u_1, u_2) \mid d(p_i, \partial D) \ge r).$$

For ϵ small enough, $C_i(u_1, u_2) \cap D = C_i(u_1, u_2)$; its area is

$$\alpha(u_1, u_2) = u_2^2 \sin^{-1} u_1.$$

If ϵ is small enough, then by Lemma 3.3, $p_j \mapsto Q(p_i, p_j)$ is a decreasing function of $|\angle(\nabla V(p_i), p_j - p_i)|$ on $C_i(u_1, u_2)$. Therefore, the probability that $p_{n(S_i)} \in C_i(u_1, u_2)$ (given that $d(p_i, \partial D) \ge r$) equals the probability that at least one node different from S_i lands in $C_i(u_1, u_2)$. Thus,

$$\begin{aligned}
\mathrm{P}(p_{n(S_i)} \in C_i(u_1, u_2) \mid d(p_i, \partial D) \ge r) &= \sum_{k=1}^{N-1} \mathrm{P}(\text{exactly } k \text{ nodes} \neq S_i \text{ lie in } C_i(u_1, u_2) \mid d(p_i, \partial D) \ge r) \\
&= \sum_{k=1}^{N-1} \binom{N-1}{k} \alpha(u_1, u_2)^k [1 - \alpha(u_1, u_2)]^{N-1-k} \\
&= 1 - [1 - \alpha(u_1, u_2)]^{N-1}.
\end{aligned}$$

Since this is true for any pair (u_1, u_2) with the above properties, it follows that

$$\begin{aligned}
\mathrm{P}(|\mathrm{Grad}(p_i) - \nabla V(p_i)| < \epsilon \mid d(p_i, \partial D) \ge r) &\ge 1 - [1 - \max_{u_1, u_2} \alpha(u_1, u_2)]^{N-1} \\
&= 1 - [1 - \mu_V(\epsilon)]^{N-1}.
\end{aligned}$$

The second part of the Proposition follows analogously. □

3.5. **Corollary.** *For every* $1 \leq i \leq N$ *and* $\epsilon > 0$,

$$\lim_{N\to\infty} \mathrm{P}(|\mathrm{Grad}(p_i) - \nabla V(p_i)| < \epsilon \mid d(p_i, \partial D) \geq r) = 1.$$

3.6. **Proposition.** *Suppose* p_i *is an equilibrium of* ∇V *and* $0 < \eta < 1$. *If*

$$N \geq N_\epsilon = 2 + \frac{\log \eta}{\log\left(1 - \frac{4\pi\epsilon^2}{H^2}\right)},$$

then

$$\mathrm{P}(|\mathrm{Grad}(p_i)| < \epsilon \mid d(p_i, \partial D) \geq r) > 1 - \eta.$$

Proof. Follows directly from Proposition 3.4. The growth of the function $\epsilon \mapsto N_\epsilon$, as $\epsilon \to 0+$ can be seen in Fig. 2. □

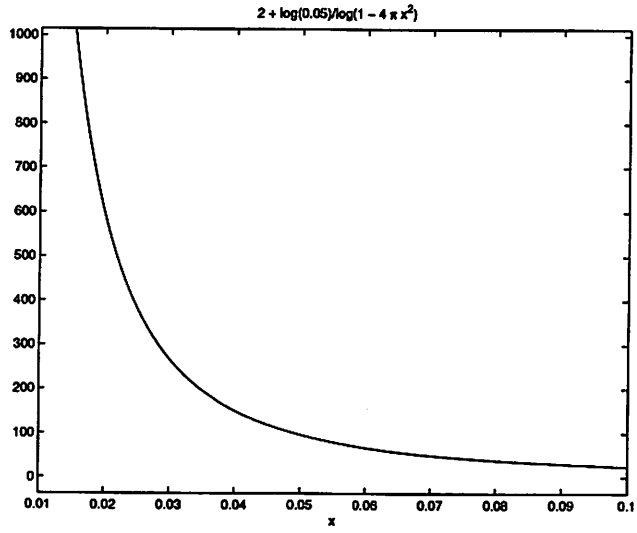

FIGURE 2. The graph of $\epsilon \mapsto N_\epsilon$ for $\eta = 0.05$ and $H = 1$.

Remark. Let $F = \nabla V$. If p is not an equilibrium of F, then in a neighborhood of p, F looks essentially like a constant vector field, up to a smooth change of coordinates. This is known as the Flow Box Theorem in dynamical systems. If $F(p) = 0$, then the picture can be much more complicated. However, if $A = DF(p)$ has no eigenvalues on the imaginary axis, then in a neighborhood of p, F looks essentially like A, or, more precisely, up to a continuous coordinate change near p, the flow of F is the same as the flow of A. This is known as the Hartman-Grobman theorem. Observe that the condition "$DF(p)$ has no eigenvalues on the imaginary axis" is generic, i.e., it is satisfied by almost all F. Furthermore, it is well known that, generically (when D^2V is nonsingular), the equilibria of ∇V can only be saddles and stable or unstable nodes.

4. Average complexity

One way to measure the average computational complexity of $GRAD_S(r)$ is to require that it satisfy an error bound such as $|\text{Grad}| < \epsilon$ with probability greater than $1-\eta$, and then count the expected number of steps the algorithm has to perform. The random variable crucial in this count is the number X_r of r-neighbors of a fixed node S_i. If the position of S_i is p_i, it is not difficult to show that

$$E(X_r \mid d(p_i, \partial D) \geq r) = (N-1)\pi r^2. \tag{2}$$

4.1. Proposition. **(a)** *If $\nabla V(p_i) = 0$, then the average computational complexity of $GRAD_S(r)$ satisfying*

$$\mathrm{P}(|\text{Grad}(p_i)| < \epsilon \mid d(p_i, \partial D) \geq r) > 1 - \eta,$$

is $O\left(\frac{1}{\epsilon}\log\frac{1}{\eta}\right)$, as $\epsilon, \eta \to 0$.

(b) *The average communication complexity of $GRAD_S(r)$ is $O(N^2)$. That is, on average, the number of messages exchanged in the execution of $GRAD_S(r)$ is of the order N^2.*

Proof. The average complexity of $GRAD_S(r)$ is $O(E(X_r \mid d(p_i, \partial D) \geq r))$, where X_r is the number of r-neighbors of S_i. Part (a) then follows from (2) Proposition 3.6. Part (b) is also a direct consequence of (2), since each of the N nodes on average has to communicate with $(N-1)\pi r^2$. □

5. Simulation results

In accordance with the previous Remark, it is sufficient to test our algorithm in three cases: near a nonequilibrium point for ∇V, near a saddle for ∇V, and near an unstable node for ∇V. Therefore, we present three examples: in the first one, V is a linear function (Fig. 3); in the second one, V is quadratic and ∇V has a saddle at $(10, 10)$ (Fig. 4); in the last one, V is quadratic, but ∇V has an unstable node at $(10, 10)$ (Fig. 5). In all cases, the algorithm gives good results away from the boundary of $D = [0, 20] \times [0, 20]$. Observe that if we excluded the edge effects from the calculation of the average relative error (i.e., average absolute error divided by the norm of the gradient at the corresponding point), the accuracy would go up.

6. Conclusion

We presented a distributed algorithm which estimates the gradient of a smooth function using a random sensor network. The method amounts to approximate differentiation of a function given its value at several random points. We estimated the probability that the error is small and showed that it converges to zero, as the number of nodes goes to infinity.

It would be useful to estimate the expected value of the error and show that it converges to zero with the number of nodes going to infinity. Furthermore, it is particularly important

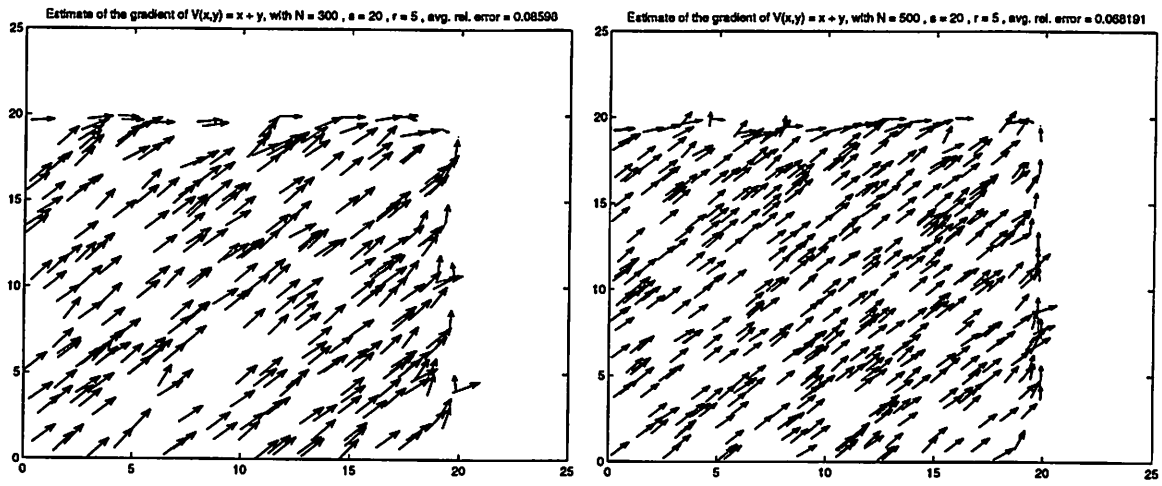

FIGURE 3. $V(x,y) = x + y$, $D = [0,20] \times [0,20]$.

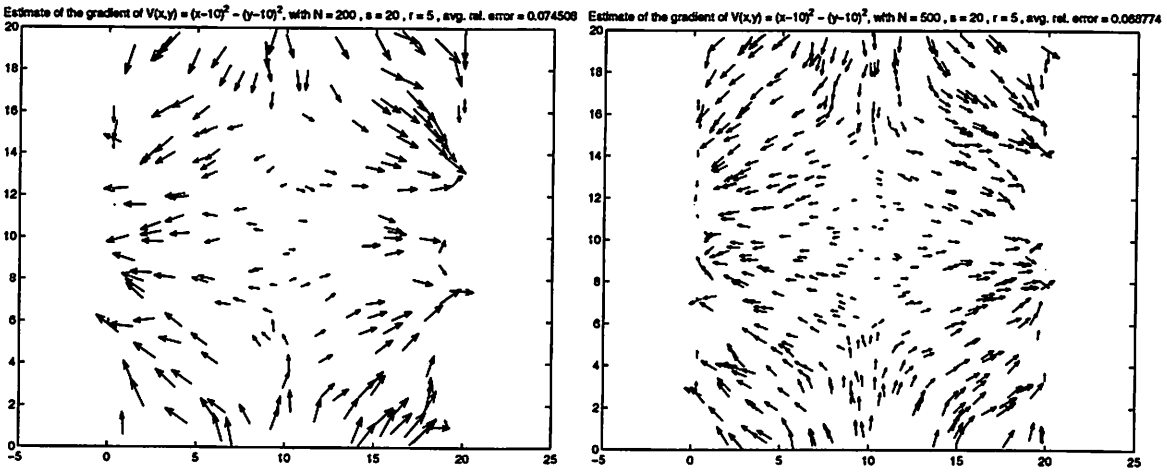

FIGURE 4. $V(x,y) = (x-10)^2 - (y-10)^2$, $D = [0,20] \times [0,20]$.

to investigate the robustness of the algorithm to noise and node failures. We plan to do this in future work.

REFERENCES

[Cul] D. Culler, `http://webs.cs.berkeley.edu`.

[Doh00] L. Doherty, *Algorithms for position and data recovery in wireless sensor networks*, Master's thesis, UC Berkeley, 2000.

[GS97] G.R. Grimmett and D.R. Strizaker, *Probability and random processes*, 2nd ed., Oxford University Press, 1997.

[KKP] J.M. Kahn, R.H. Katz, and K.S.J. Pister, *Mobile networking for Smart Dust*, ACM/IEEE Intl. Conf. on Mobile Computing and Networking (Seattle, WA), August 1999.

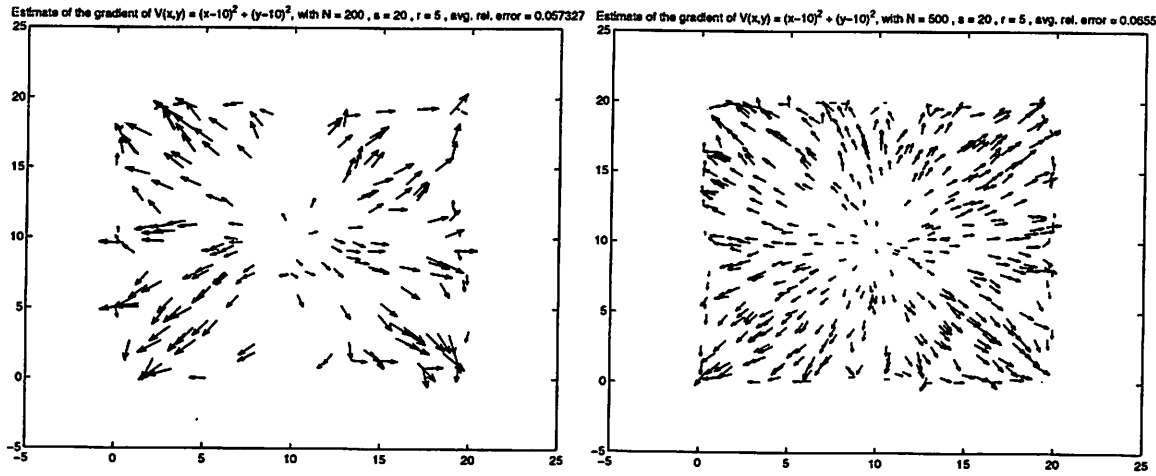

FIGURE 5. $V(x,y) = (x-10)^2 + (y-10)^2$, $D = [0,20] \times [0,20]$.

[SS01] S. N. Simić and S. Sastry, *Distributed localization in wireless ad hoc networks*, December 2001, preprint available at `http://robotics.eecs.berkeley.edu/~simic/PDF/localization.pdf`.

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCES, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720

E-mail address: `simic@eecs.berkeley.edu`, `sastry@eecs.berkeley.edu`