Distributed Environmental Monitoring Using Random Sensor Networks***

Slobodan N. Simić and Shankar Sastry

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720

{simic,sastry}@eecs.berkeley.edu

Abstract. We present a distributed algorithm for environmental monitoring of a scalar field (such as temperature, intensity of light, atmospheric pressure, etc.) using a random sensor network. We derive an error estimate, discuss the average complexity of the algorithm, and present some simulation results.

Introduction

In this paper, we present a distributed algorithm for estimating the gradient of a scalar field (such as temperature, intensity of light, atmospheric pressure, etc.) using a random wireless sensor network. Environmental monitoring is one of the main applications of the emerging technology of wireless sensor networks. Our algorithm has potential applications in preventing forest fires, energy conservation, oceanography, building science, etc. We envision using a large number of sensor nodes to automatically detect the emergence of critical points (such as heat sources in the context of forest fires) and notifying the base station which can then take further action.

This 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 theoretical work on environmental monitoring using random sensor networks was done in [Doh00]. For a study of wireless sensor networks in real-world habitat monitoring, see [MPS+02]. We also mention [MEM01], which deals with gradient estimation from scattered data in geology.

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

** *Proceedings of the* 2^{nd} *International Workshop on Information Processing in Sensor Networks*, PARC, Palo Alto, 2003, LNCS 2634, pp. 582-592.

Due to high long range communication costs, low battery power, and need for robustness to node failures, 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, when necessary, be relayed to a base station.

The basic idea of our algorithm is the following. Each node communicates with its neighbors and computes the maximal difference quotient of the sensed scalar field. 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 not new or sophisticated; however it has two redeeming features: (1) it is sufficiently simple and computationally non-intensive to be implementable on the current sensor network platform [Cul]; (2) it allows us to compute analytic error estimates. We are able to *prove* that, in a probabilistic sense, the algorithm converges (i.e., as the number of nodes goes to infinity, the probability that the error is as small as we want converges to one), and to answer questions like "What should the number of nodes be so that the probability that the error is less than some ϵ, is greater than $1 - \eta$?"

We believe that in the sensor network literature, there is a need for a more precise theoretical analysis of known problems and proposed solutions. We therefore emphasize that the main purpose of this paper is to rigorously analyze the accuracy and complexity of the proposed algorithm from a probabilistic point of view, *not* to discuss any implementation issues, which will be dealt with in future work.

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 error estimates; Section 4 discusses average complexity, followed by 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 $S_1, \ldots, S_N$ is deployed in some region $D \subset \mathbb{R}^2$. The number i will be called the ID of the node S_i. We make the following simplifying assumptions:

– 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 [SS02]).
– 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. In this case, two nodes whose distance is $\leq r$ are called r-*neighbors*.

Our *goal* is:

> Using only local information, design a distributed algorithm for estimating the gradient of V at $p_1, \ldots, p_N$, and find its error.

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 $\mathbb{R}^2$.
– Random variables $p_1, \ldots, p_N$ are independent and uniformly distributed in 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 *COLLECT IDs, positions, and measurements from all r-neighbors.*
Each r-neighbor S_ν contributes (ν, p_ν, v_ν), 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* S_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 data have been processed. The* ***estimate*** *of* $\nabla V(p)$ *is*

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

Note that $v_\nu = V(p_\nu)$; $q(S)$ is the current value of the maximal difference quotient, and $n(S)$ is the ID of the corresponding node.

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

Presently, we assume that in Step 2 we use one of the existing *data fusion* algorithms. We refer the reader to some of the relevant data fusion literature such as [KM94,QWIC01,IJ01,GDW94]. We are currently investigating this problem in the context of environmental monitoring, but for reasons of space, we postpone its discussion to future work.

3 Error estimates for $\boldsymbol{GRAD_S(r)}$

In this section we estimate the error of the proposed algorithm. The proofs of all statements are elementary and are therefore included, but can be skipped in first reading.

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

Proposition 1 *For all* $p, q \in D$, $p \neq q$,

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

Proof. By the Fundamental Theorem of Calculus,

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

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

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

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

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

Thus,

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

It is not hard to see that

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

This completes the proof of the Proposition. ■

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

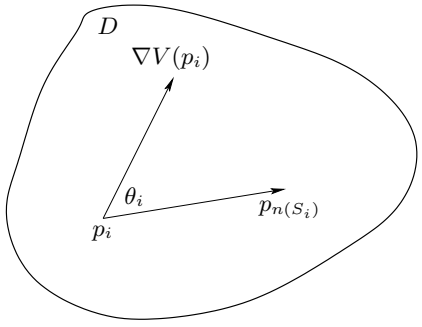

Fig. 1. The angle θ_i.

Corollary 1 *For every* $1 \le i \le N$,

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

The following lemma says that if we are sufficiently close to a node, it is the direction that matters in estimating the gradient.

Lemma 1 *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, $q_i \mapsto Q(q, q_i)$ *is a decreasing function of* α_i.

Therefore, if q_i's are sufficiently close to q and the angles $\angle(\nabla V(q), q_i - q)$ are not too big, then the difference quotient $Q(q, q_i)$ increases as the vector $q_i - q$ becomes more parallel to $\nabla V(q)$.

Proof. Let $A = |\nabla V(q)|$ and

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

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

$$\rho < \min\left\{\frac{Ac}{2H} : 1 \le i \le K\right\}.$$

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, $\mathrm{I} \ge Ac$ and $|\mathrm{II}| \le 2\rho H$. Therefore,

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

implying $Q(q, q_i) > Q(q, q_j)$. ∎

Denote by $\mathbb{P}(A|B)$ and $\mathbb{E}(A|B)$ 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. Also let

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

Proposition 2 *For all $1 \leq i \leq N$ and $\epsilon > 0$ small enough,*

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

where

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

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

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

Proof. Let $C_i(u_1, u_2)$ (Fig. 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 $\nabla V(p_i)$. Assume $A_i u_1 + \frac{1}{2} H u_2 < \epsilon$ and $d(p_i, \partial D) \geq r$. If $p_{n(S_i)}$, the node which realizes the maximal difference quotient among the neighbors of S_i, belongs to $C_i(u_1, u_2)$, then by Corollary 1, $|\mathrm{Grad}(p_i) - \nabla V(p_i)| \leq A_i u_1 + \frac{1}{2} H u_2 < \epsilon$. Therefore,

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

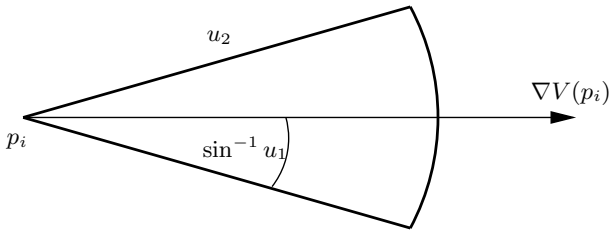

Fig. 2. The circular sector $C_i(u_1, u_2)$.

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 sufficiently small, then by Lemma 1, $p_j \mapsto Q(p_i, p_j)$ is a decreasing function of $|\angle(\nabla V(p_i), p_j - p_i)|$ on $C_i(u_1, u_2)$. Therefore, if at least one node is in

$C_i(u_1, u_2)$, then $p_{n(S_i)} \in C_i(u_1, u_2)$; the converse is clear enough. Hence the probability that $p_{n(S_i)} \in C_i(u_1, u_2)$ (given $d(p_i, \partial D) \geq r$) equals the probability that at least one node different from S_i lands in $C_i(u_1, u_2)$, so

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

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

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

If p_i is an equilibrium of ∇V, then $A_i = 0$. By Corollary 1, $|\text{Grad}(p_i)| < \epsilon$ if $|p_{n(S_i)} - p_i| < 2\epsilon/H$, so $\mathbb{P}(|\text{Grad}(p_i)| < \epsilon |\ d(p_i, \partial D) \geq r)$ is not less than the area of the disk centered at p_i of radius $2\epsilon/H$, proving the second part of the Proposition. ■

Corollary 2 *For every $1 \leq i \leq N$ and $\epsilon > 0$,*

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

Therefore, the algorithm, in the sense specified by the Corollary, "converges in probability".

Proposition 3 *Suppose p_i is an equilibrium of ∇V and $0 < \eta < 1$. If*

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

then

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

Proof. Follows directly from Proposition 2. Observe that as $\epsilon, \eta \to 0$, $N(\epsilon, \eta)$ is of the order $O\left(\frac{1}{\epsilon} \log \frac{1}{\eta}\right)$.

4 Average complexity

One way to measure the average complexity of $GRAD_S(r)$ is to require that the probability that $|\text{Grad}| < \epsilon$ be greater than $1 - \eta$, and then count the average

number of computations and communication steps the algorithm has to perform. The random variable crucial in this count is the number X_r of r-neighbors of a randomly picked but fixed node S_i. If the position of S_i is p_i, it is not difficult to show that

$$\mathbb{E}(X_r \mid d(p_i, \partial D) \geq r) = (N-1)\pi r^2. \tag{3}$$

Proposition 4 *If $\nabla V(p_i) = 0$ and $N \geq N(\epsilon, \eta)$, guaranteeing that*

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

than on average, the number of computations $GRAD_S(r)$ performs is of the order $O\left(\frac{1}{\epsilon}\log\frac{1}{\eta}\right)$, as $\epsilon, \eta \to 0$.

Proof. The number of computational steps S performs in $GRAD_S(r)$ is proportional to the number of its r-neighbors, that is, on average, of the order $O(E(X_r \mid d(p_i, \partial D) \geq r)) = O(N)$. The statement then follows from (3) and Proposition 3, since N has to be chosen of the order $O\left(\frac{1}{\epsilon}\log\frac{1}{\eta}\right)$.

Remark. The average communication complexity of the algorithm depends on the data fusion algorithm chosen in Step 2.

5 Simulation results

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.

Therefore, 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. Consequently, 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 improve.

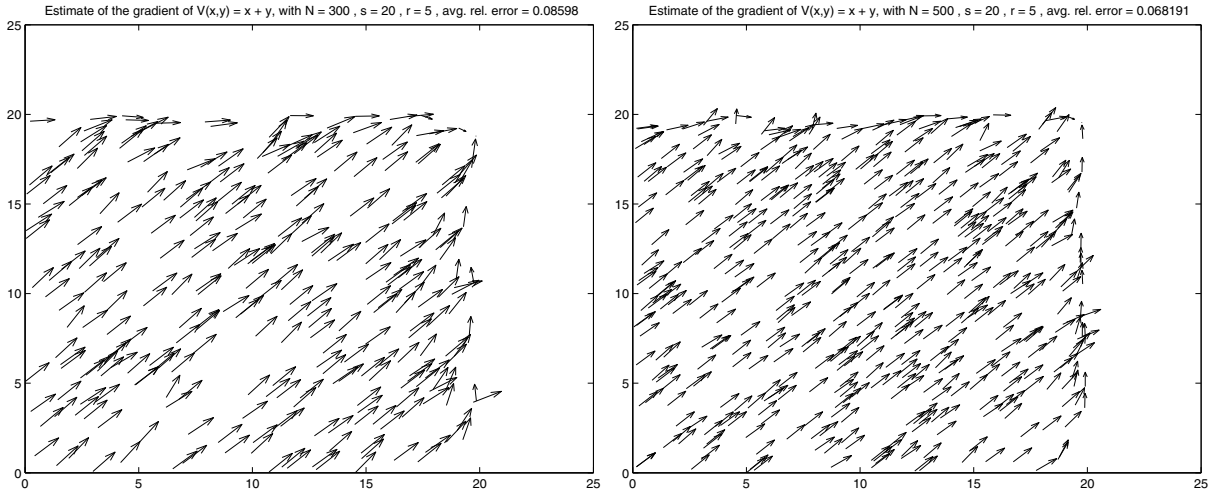

Fig. 3. $V(x, y) = x + y$, $D = [0, 20] \times [0, 20]$.

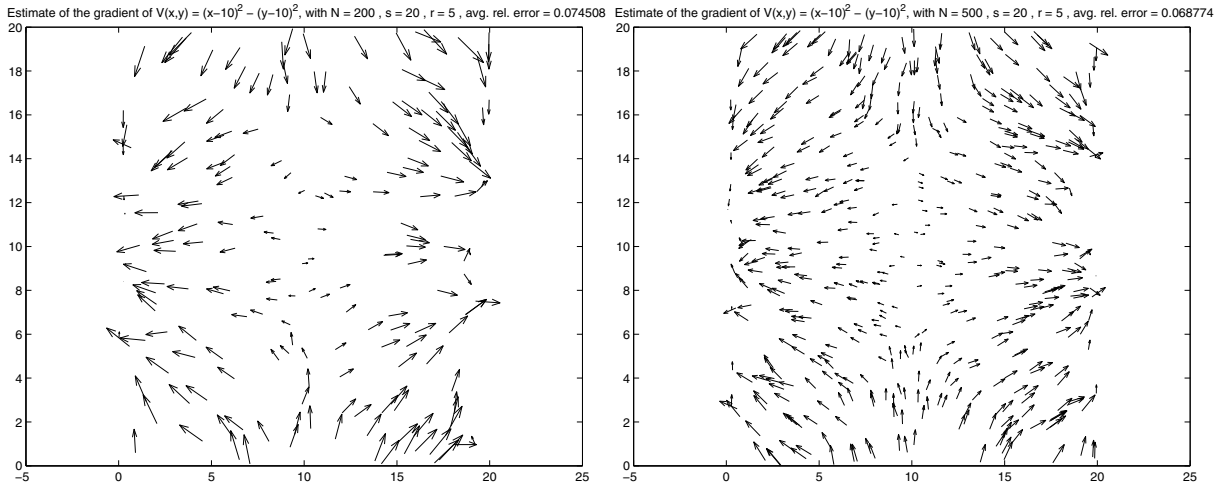

Fig. 4. $V(x, y) = (x - 10)^2 - (y - 10)^2$, $D = [0, 20] \times [0, 20]$.

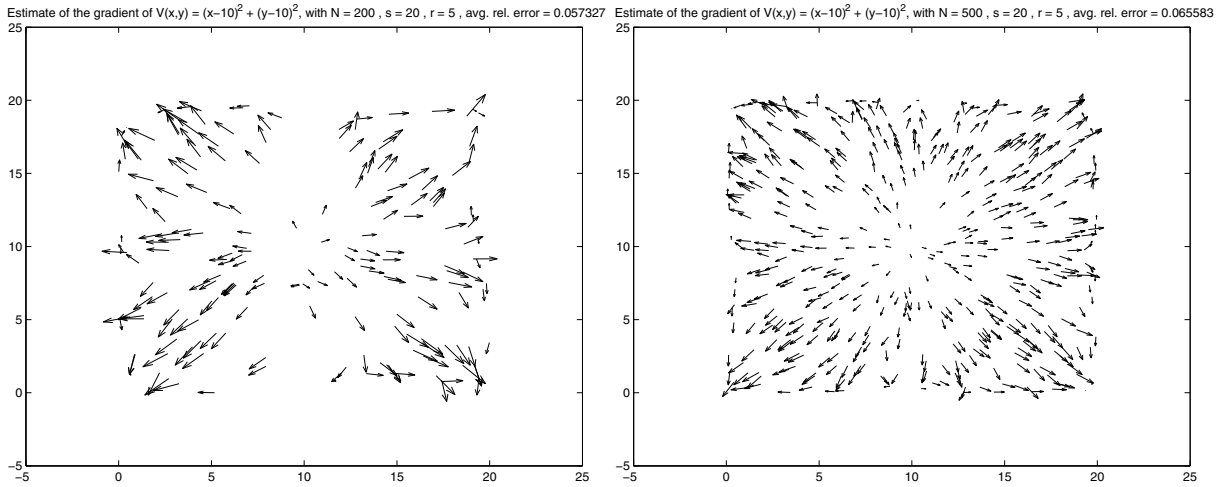

Fig. 5. $V(x, y) = (x - 10)^2 + (y - 10)^2$, $D = [0, 20] \times [0, 20]$.

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 the function given its value on a set of random points. We estimated the probability that the error is small and showed that it converges to one, as the number of nodes goes to infinity.

It would be useful to estimate the expected value of the error and investigate robustness of the algorithm to noise and node failures. Further, it would be interesting to compare this algorithm with others, e.g., the one in [MEM01], which is also sufficiently simple to be implementable on the current platform for sensor network. 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.

[GDW94] S. Grime and H. F. Durrant-Whyte, *Data fusion in decentralized sensor networks*, Control Enineering Practice **2** (1994), no. 5, 849–863.

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

[IJ01] S. S. Iyengar and B. Jones, *Information fusion techniques for pattern analysis in large sensor data networks*, Journal of the Franklin Institute **338** (2001), no. 5, 571–582.

[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.

[KM94] A. Knoll and J. Meinkoehn, *Data fusion using large multi-agent networks: an analysis of network structure and performance*, Proceedings of the International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI) (Las Vegas, NV), Oct. 2-5 1994, pp. 113–120.

[MEM01] T. H. Meyer, M. Eriksson, and R. C. Maggio, *Gradient estimation from irregularly spaced data sets*, Mathematical Geology **33** (2001), no. 6, 693–717.

[MPS+02] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, *Wireless sensor networks for habitat monitoring.*, ACM International Workshop on Wireless Sensor Networks and Applications, WSNA '02 (Atlanta, GA), September 2002.

[QWIC01] H. Qi, X. Wang, S. S. Iyengar, and K. Chakrabarty, *Multisensor data fusion in distributed sensor networks using mobile agents*, Proc. Intl. Conf. Information Fusion, August 2001, pp. 11–16,.

[SS02] S. N. Simić and S. Sastry, *Distributed localization in wireless ad hoc networks*, Tech. report, UC Berkeley, 2002, Memorandum No. UCB/ERL M02/26.