AIAA Guidance, Navigation and Control Conference and Exhibit
20 - 23 August 2007, Hilton Head, South Carolina

AIAA 2007-6751

Decentralized Estimation under Communication Constraints

Emily M. Craparo,* Jonathan P. How,†
Daniela Pucci de Farias‡ and Nicholas Roy§
Massachusetts Institute of Technology

The decentralized cooperative exploration problem necessarily involves communication among agents, while the spatial separation inherent in this task places fundamental limits on the amount of data that can be transmitted. However, the impact of limited communication on the exploration process has not been fully characterized. No known exploration algorithm realistically models the tradeoff between rapid expansion (which allows more rapid exploration of the map) and maintaining close relative proximity among agents (which facilitates communication). This work is a first step toward characterizing the impact of limited communication on this cooperative estimation task by considering a static version of the problem. The mathematical properties of the information form of the Kalman filter are leveraged in the development of two approximate algorithms for selecting highly informative portions of the information matrix for transmission. One algorithm, a fully polynomial time approximation scheme, provides provably good results in computationally tractable time for problem instances of a particular structure. The other, a heuristic-based method applicable to instances of arbitrary matrix structure, performs very well in simulation for randomly-generated problems of realistic dimension.

I. Introduction

One of the fundamental issues facing autonomous mobile robots is the necessity of operating in *a priori* unknown environments. The ability to autonomously build a map of an environment and, if necessary, estimate the robot's own location within that map is regarded as a primary prerequisite for fully autonomous robots. This challenge is captured by the well-known simultaneous localization and mapping (SLAM) problem.[20]

Scalability is an important characteristic in algorithms for collaborative mobile robot teams.[19] The magnitude of problem parameters and the number of robots available to perform tasks are both unknown *a priori* in many problems of interest.[10] Without the ability to adapt a solution structure to problem instances of practical scale, the full benefit of using multiple robots is not realized. The need to create cooperative algorithms that scale well with the size of a robot team and with the magnitude of problem parameters (e.g. the number of features in a mapping problem) has led to a widespread interest in decentralized and approximate algorithms for many problems.[3,4,12]

In feature-based exploration and mapping problems, true scalability implies the ability of a large number of robotic agents to cooperatively estimate the location of a large number of features

*Research Assistant; MIT Dept. of Aeronautics and Astronautics; AIAA Student Member.
†Professor; MIT Dept. of Aeronautics and Astronautics; AIAA Associate Fellow.
‡Assistant professor; MIT Dept. of Mechanical Engineering;
§Assistant professor; MIT Dept. of Aeronautics and Astronautics; AIAA Member.

spread over a large area. Large numbers of agents generally rely on decentralized algorithms in which communication among agents facilitates the coordination of planning.[1] However, when the number of features is very large, it is impossible for all agents to communicate all information to all other agents. Furthermore, because of the spatially distributed nature of the exploration task, communication may be hindered by large separations between agents. The effects of path loss, which increase with separation distance,[16] result in decreasing data rate for a fixed transmission power as distance between agents increases. Thus, it is necessary that decentralized algorithms for autonomous mapping account for the fact that not all information can be transmitted between agents, and that agents must be able to act in situations where only partial information is available.

II. Background

The exploration problem has been examined, in various forms, for many years. Some early work was done by Smith, Self and Cheeseman,[18] and Leonard,[13] Durrant-Whyte and Thrun[22] continue to be major contributors to the field. Significant work has also been done on the mapping problem; that is, the problem of deriving a probabilistic representation of the world from sensor measurements and other observations. A particular challenge in this area is the problem of merging maps from various sources, as occurs in the decentralized exploration problem. Work in this area has been done by Konolige and Fox.[9] Also of relevance are techniques for combining estimates in provably consistent ways, such as covariance intersection[24] and various modifications to this technique.

While the mapping problem has been studied in great detail, and in fact decentralized algorithms exist for performing mapping,[27,28] none of these techniques adequately address the issue of communication among agents or capture the tradeoff between agents expanding their separation (in order to examine the map more quickly) and staying close to one another (in order to maintain adequate communication links). Some approaches in this direction have involved, for example, constraining the amount of information that can be sent between agents in any one time step. This approach does not allow agents to take advantage of close proximity to send large amounts of data, nor does it realistically model the cost of communicating over very large distances. Other approaches have involved setting a maximum inter-agent distance in order to maintain connectivity, but this approach does not balance the reward of information gained by increasing inter-agent separation with the increased cost of communicating: the inter-agent separation is set *a priori* and does not depend on specific characteristics of the map, or the status of the exploration process. Thus, our eventual aim is to develop an algorithm that explicitly balances these competing objectives.

Work has been done by Nettleton *et al.*[15] in the area of decentralized estimation under communication constraints. In the algorithm presented in their work, the information to be sent from the sender's information matrix Y is selected as the single submatrix containing the features about which the sender has learned the most since the previous transmission. This is a reasonable approach, but it may be improved upon. This will be discussed further in Section B.

The work presented in this paper leverages the properties of the information filter, which is mathematically equivalent to the traditional Kalman filter but which possesses properties of value for the decentralized exploration problem.[6,14] The information matrix and information vector are defined as $Y = P^{-1}$ and $y = P^{-1}\hat{x}$, where P is the covariance matrix in the traditional Kalman filter, and $\hat{x}$ is the state vector. Compared to the covariance filter, the information filter essentially exchanges complexity in the update step for complexity in the prediction step.

Denoting the state transition matrix by F, the observation matrix by H, the process noise covariance matrix by Q, and the observation noise covariance matrix by R, the prediction step of the information filter can be written as

$$
\begin{aligned}
Y(k|k-1) &= [F(k)Y^{-1}(k-1|k-1)F^T(k) + Q(k)]^{-1} && (1)\\
y(k|k-1) &= Y(k|k-1)F(k)Y^{-1}(k-1|k-1)y(k-1|k-1) && (2)
\end{aligned}
$$

and the measurement update step is:

$$
\begin{aligned}
Y(k|k) &= Y(k|k-1) + H^T(k)R^{-1}(k)H(k) && (3)\\
y(k|k) &= y(k|k-1) + H^T(k)R^{-1}(k)z(k) && (4)
\end{aligned}
$$

The information filter is particularly well suited for decentralized sensor fusion problems in which many measurement updates may take place in a single time step.[6,14] This additive structure also facilitates the development of the algorithms described in this paper.

III. Model

The act of exploring an unknown environment involves agents moving, sensing the locations of features, and communicating with one another (or with a central repository) about the locations of features. In the case of decentralized exploration, it is also beneficial for agents to exchange information about their planned future actions.[1] All of these activities require energy, a resource that is often tightly constrained for small mobile robots. The relatively small robotic agents that tend to be utilized most often for decentralized schemes are the very ones that are most challenged in terms of power. Thus, expected energy expenditure provides a reasonable cost metric for potential plans under consideration, and in this work all costs in the mapping problem are modeled in terms of energy expenditure.

Exploration is, in essence, an information-gathering process. If a probabilistic representation of the map estimate is assumed,[21] the quality of a map estimate can be expressed as its Shannon entropy, or the degree to which the probability distribution of feature location estimates is compact or spread. In the case of a Gaussian distribution, the entropy of the distribution is also related to the mean squared error of the estimate, a quantity of importance in control applications. Therefore, entropy is the quality metric used for this work.

The centralized version of the mapping problem can be expressed as a minimization of entropy, subject to an energy constraint:

$$
\begin{aligned}
&\min_{m_{it}, s_{it}, c_{it}} && U_T(m_{it}, s_{it}, c_{it}) && (5)\\
&\text{subject to} && \sum_{i=1}^{n}\sum_{t=0}^{T} E(m_{it}, s_{it}, c_{it}) \le E_{max} && (6)
\end{aligned}
$$

where m_{it} denotes agent i's motion at time t, s_{it} denotes agent i's sensing at time t, c_{it} denotes agent i's communication to other agents or to a central repository at time t, U_T denotes entropy at a central repository at the final time T, and E denotes energy. Note that although total energy is minimized in this formulation, other formulations are also of interest, including minimization of the maximum energy expenditure by any agent. An alternative problem formulation involves minimizing energy expenditure subject to an entropy constraint:

$$
\begin{aligned}
&\min_{m_{it}, s_{it}, c_{it}} && \sum_{i=1}^{n}\sum_{t=0}^{T} E(m_{it}, s_{it}, c_{it}) && (7)\\
&\text{subject to} && U_T(m_{it}, s_{it}, c_{it}) \le U_{max} && (8)
\end{aligned}
$$

It is desirable that a solution to this problem scale well with the number of map features, the spatial dimension of the map, and the number of agents performing the exploration. However, it rapidly becomes intractable to solve this problem to optimality for large numbers of agents, and communication limitations make it impractical for a central decision-maker to maintain adequate communication with all agents across large distances. In order to ensure scalability it is therefore appropriate to consider a decentralized version of the exploration problem, in which decision making takes place at an agent level rather than through a central decision-maker. In this case, an appropriate figure of merit is the average entropy of the estimates of all agents in a team rather than the entropy of the estimate at a central repository, and each agent is subject to its own energy constraint.

IV. The Static Problem

A first step toward developing a sequential decision-making algorithm is understanding the implications of a single decision. In the case of deciding the amount of energy to expend making communication transmissions in the exploration problem, one must know (or have an estimate of) the additional information that can be gained for a given amount of information transmitted before deciding whether it is worthwhile to transmit this amount of information. This prompts the question: what is the maximum amount of information that can be gained by transmitting a quantity of information not more than bandwidth B?

To answer this question, inter-agent communication problem is examined in a static sense. As discussed above, the exploration problem is a multi-step process in which agents must decide their motion, sensing, and communication plans sequentially at every time step. A challenging subproblem is simply to decide, for a single pair of agents in a single time step, what information should be communicated between them in order to minimize the entropy of the agents' probability distributions after incorporating the communicated information. Given a solution, or approximate solution, to this problem, one can then construct a dynamic planning algorithm which uses this greedy communication strategy. While the greedy strategy does not necessarily provide the best solution to a multi-step problem, it does provide valuable insights into the nature of the utility of information in the decentralized estimation problem.

In the problem under consideration in this paper, two agents jointly estimate the location of a number of stationary features. Without loss of generality, one agent is designated as the sender and the other the receiver. Both agents represent their feature estimates as Gaussian distributions, in the inverse covariance form. It is assumed that the sender has a perfect estimate of the receiver's information matrix, and the goal is simply to choose the most beneficial elements of the receiver's own information matrix to send to the receiver. The utility of information is evaluated in terms of the reduction in entropy of the estimate, which in the Gaussian case is represented by

$$
\begin{aligned}
U &= \frac{1}{2}\log[(2\pi e)^n|\det(P)|] = \frac{1}{2}[\log((2\pi e)^n) + \log(|\det(P)|)] && (9)\\
&= \frac{1}{2}[\log((2\pi e)^n) + \log(\left|\det(Y^{-1})\right|)] = \frac{1}{2}[\log((2\pi e)^n) + \log(\frac{1}{|\det(Y)|})] && (10)\\
&= \frac{1}{2}[\log((2\pi e)^n) - \log(|\det(Y)|)] && (11)
\end{aligned}
$$

where P is the receiver's covariance matrix, and Y is the information matrix, after the receiver has incorporated the information transmitted by the sender.

It is assumed that it is not possible for the agents to exchange an arbitrarily large amount of information. In particular, the sender has a limit on the amount of information that can be sent, denoted by bandwidth B. For simplicity of modeling, assume that each element of the information matrix requires a single unit of bandwidth to send. There is also a certain amount of communication overhead required to provide the receiver with labels describing which diagonal elements are being sent, and this cost is modeled as one unit of bandwidth per diagonal element. (It is assumed and both agents are estimating the location of a common set of features, and that they share a system for labeling them.) Note that for a given set of diagonal elements to be communicated, the sender must choose whether to send information about off-diagonals, compounding the difficulty of the problem. Given the entropy metric and the bandwidth cost model, if the covariance intersection algorithm is used to combine the receiver's information matrix with the transmitted submatrices, the sender's problem is

$$
\begin{aligned}
\text{minimize} \quad & c^T x + \log(\det(Y(x)^{-1})) \\
\text{subject to} \quad & Y(x) > 0 \\
& Y(x) = \sum_{i=0}^{M} x_i y_i Y_i \\
& \sum_{i=1}^{M} y_i B_i \leq B \\
& y_i \in \{0,1\} \quad\quad i = 1, ..., M \\
& y_0 = 1 \\
& 0 \leq x_i \leq 1 \quad\quad i = 0, ..., M \\
& \sum_{i=0}^{M} x_i = 1
\end{aligned} \tag{12}
$$

where Y_0 is the receiver's initial information matrix; Y_i is the i^{th} candidate submatrix, extracted from the original information matrix as described in [15]; y_i is a binary decision variable indicating that the i^{th} submatrix is selected for transmission; x_i is the weighting coefficient used in the covariance intersection algorithm; B_i is the bandwidth requirement of sending submatrix Y_i; N is the number of map features; and M is the number of submatrices of the information matrix.

This problem rapidly becomes difficult as the number of features to be mapped increases. For an $N \times N$ information matrix, the number of possible combinations of diagonal elements that could be sent is 2^N. Given a selection of diagonal elements, there is also a decision to be made regarding which cross-correlation terms should be sent. Thus, the optimal combination of data to send cannot be found through brute force search. A more intelligent strategy is needed, and we turn our attention to this topic.

A. Block-Diagonal Case

We first consider the case of a natural measurement scenario that results in a structured information matrix that is not fully populated. Assume that a mobile robot, the location of which is adequately known (through GPS or other means), is taking measurements of feature locations through an imprecise means. In this example, bearing-only measurements of feature locations are taken by means of a camera whose exact angle relative to the robot is not known, but is only estimated to within a few degrees.

If the locations of the features are being estimated in a two-dimensional plane, the x and y locations of each feature are correlated through the uncertainty in the camera angle when measurements of the feature are taken. Because the robot's location is known, however, there is no correlation from feature to feature. Thus, the information matrix is block-diagonal with 2×2 blocks:

$$
Y = \begin{bmatrix} a_1 & c_1 & 0 & 0 & \\ c_1 & b_1 & 0 & 0 & \\ 0 & 0 & a_2 & c_2 & \dots \\ 0 & 0 & c_2 & b_2 & \\ & & \vdots & & \ddots \end{bmatrix} \tag{13}
$$

Recall that the term in the entropy equation that can be altered is $\log(\det(P))$, which is to be minimized, or equivalently $\log(\det(Y))$, which is to be maximized. Because there are exponentially many possible values for Y, it is typically not feasible to solve this problem optimally, nor is there an obvious approximation algorithm for this problem. For the case of a block diagonal matrix, however, the objective decomposes into a sum of terms:

$$
\log(\det(Y)) = \log\left(\begin{vmatrix} a_1 & c_1 \\ c_1 & b_1 \end{vmatrix}\right) + \log\left(\begin{vmatrix} a_2 & c_2 \\ c_2 & b_2 \end{vmatrix}\right) + \ldots + \log\left(\begin{vmatrix} a_m & c_m \\ c_m & b_m \end{vmatrix}\right) \tag{14}
$$

For each block in this matrix, there are five possible decisions that can be made by the sender. The sender could send:

1. Information about a only,
2. Information about b only,
3. Information about a and b with no off-diagonal information,
4. Information about a and b coordinates with cross correlation information, or
5. Nothing at all.

Under our communication model, the possible transmissions would require a bandwidth of 2, 2, 4, 5, and 0 units, respectively, and they would result in estimated reductions in entropy that are computable given an estimate of the receiver's information matrix. A brute force approach to solving this problem would still involve searching over $O(5^m)$ combinations of transmissions. However, because the objective function is now additive, the problem takes on the form of a multiple-choice knapsack problem.

The multiple-choice knapsack problem is a variant of the traditional knapsack problem in which each item belongs to one of several disjoint classes, and the goal is to choose exactly one item from each class such that the total profit of these items is maximized. This problem is NP-hard, but there exist fully polynomial-time approximation schemes (FPTASs) for solving it.[8] In the case of the map communication problem there is an item class for each feature, and the items correspond to specific decisions that may be made by the sender. The item weights are represented by the bandwidth required to send the information (including zero if no information is communicated about a given feature), and the values of the items are represented by the expected reduction in the receiver's entropy resulting from their communication.

Using a dynamic programming-based FPTAS for this problem based on that of Lawler,[11] a performance guarantee of $(1-\epsilon)$ can be achieved with a running time of $O(\frac{ml}{\epsilon})$, where m is the number of item classes and l is the number of items per class. That is, if the value of the optimal solution to this problem is E^* and the value of the solution generated by our algorithm is $\tilde{E}$, then it is guaranteed to be the case that

$$
(1-\epsilon)E^* \leq \tilde{E} \leq E^* \tag{15}
$$

if the profits involved are integers, with a straightforward modification necessary in the case of non-integer profits. In order to evaluate the performance of this algorithm, it is compared to a simple greedy strategy. In the greedy strategy it is assumed that only complete feature submatrices will be transmitted, i.e. 2×2 blocks that lie along the diagonal of the sender's information matrix. In order to choose blocks to send, the difference in magnitude between the diagonal elements of the sender's information matrix and the receiver's information matrix is used as a ranking heuristic, and the block with the greatest total difference is ranked highest.

The performance of these two algorithms for a fixed communication bandwidth ($B = 25$) and various numbers of features m , averaged over 20 randomly generated simulation runs per data point, is shown in Figure 1. In these simulations, the locations of m features are randomly generated in a plane, and two agents take a predetermined series of measurements of them. These measurements have a known uncertainty associated with them, which allows the computation of an information matrix for each agent.

The heuristic greedy algorithm is much faster than the approximation algorithm, and for small numbers of features, it compares fairly well in terms of performance. (Note that for $B = 25$, all data may be sent for very small numbers of features presented here.) However, as the number of features increases and a trivial solution is no longer available, the approximation algorithm surpasses the greedy algorithm in terms of performance, while the difference in computation time between them remains small. Of particular note is the variability in solution quality from trial to trial; the small difference in solution quality for the approximation algorithm contrasts greatly with the high level of variability in solution quality for the heuristic.

Figure 2 compares the performance characteristics of the two algorithms for a fixed number of features (20) and an increasing availability of communication bandwidth. In this case, the approximation algorithm makes better use of increasing availability of bandwidth, with little increase in computation time. Again, the approximation algorithm exhibits great consistency in solution quality due to its performance guarantee, while the quality of the heuristic solution is highly variable.

B. Fully-Populated Case

The measurement scheme described in the previous section assumed perfect knowledge of vehicle location in order to arrive at a block diagonal form for the information matrix. In a more general scenario, the vehicle location might be unknown and would have to be estimated along with feature locations using the same corrupted sensor measurements. This simultaneous localization and mapping (SLAM) problem naturally leads to a fully-populated information matrix,[26] although many terms are often small in the normalized information matrix. There exist sparsification techniques that seek to take advantage of the small magnitude of these values in the normalized information matrix to reach an approximate matrix which is sparse. However, some of these sparsification techniques produce solutions that overestimate confidence in feature locations.[26] Other sparsification techniques work by managing the number of terms of cross correlation and do not explicitly manage the size of blocks in the information matrix. Note that in the previous section, the total number of items in the multidimensional knapsack problem was $5m$ because for each 2x2 block, only five possible choices could be made. If the block size were to increase, however, the number of choices that could be made would grow exponentially with the dimension of the block. Thus, the solution technique applied in the previous section does not scale well as the size of the blocks in the information matrix grows.

The case of the fully-populated information matrix has been examined in [15], in which a single large block from the information matrix containing all diagonal and off-diagonal information is transmitted. The features selected for transmission are those about which the sender has learned the most since the previous transmission. However, because of the constraint that only a single large submatrix is sent, a great deal of bandwidth is used to transmit relatively unimportant cross correlation terms in many cases.[26]

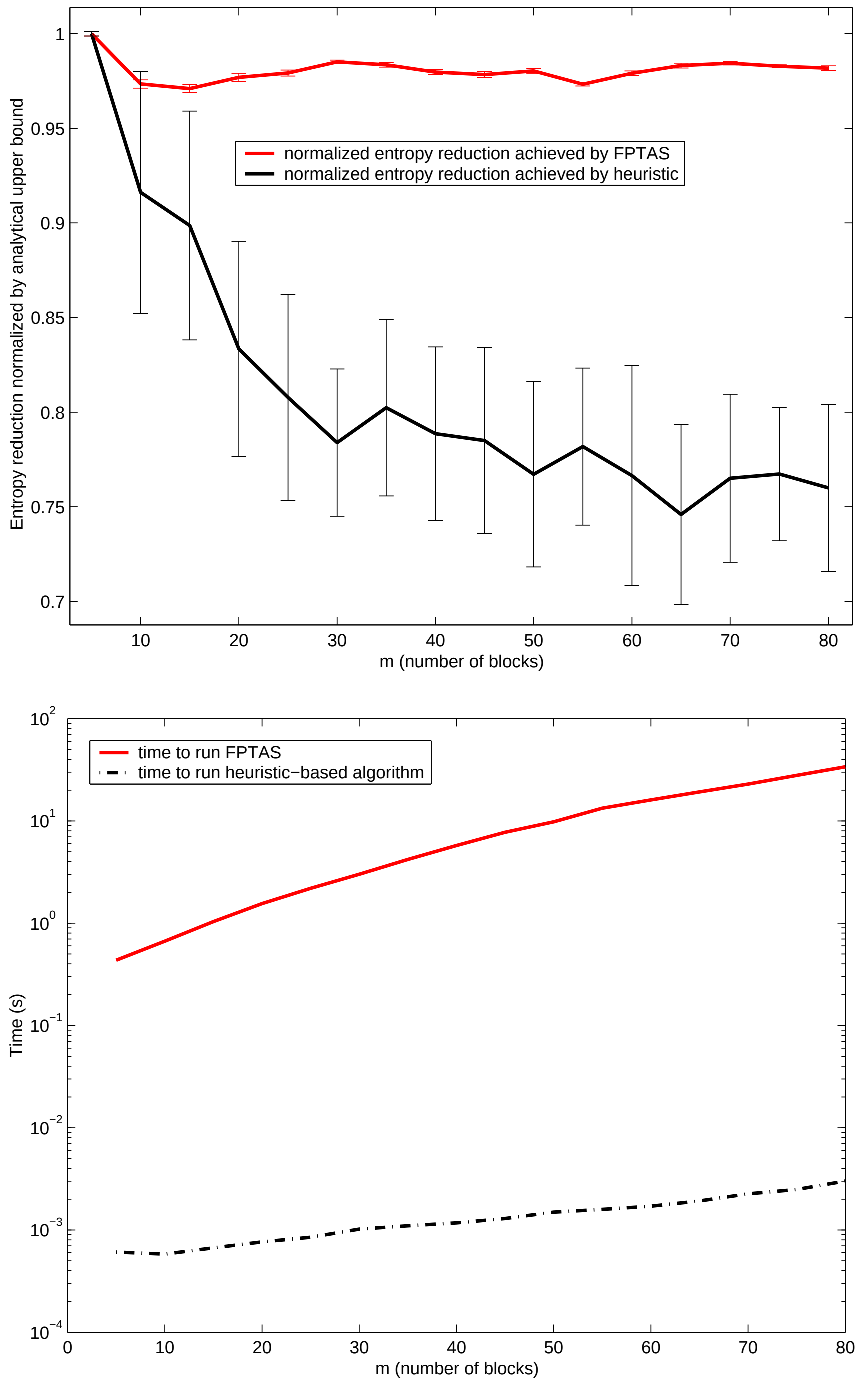

Figure 1. The performance of a simple greedy algorithm is compared to that of the fully polynomial time approximation algorithm developed in this section for various numbers of features and a fixed communication bandwidth of 25 units. The top figure depicts the reduction in the entropy of the receiver's estimate achieved by each algorithm, normalized by the maximum possible reduction in entropy as given by the performance bound of the approximation algorithm. Each data point represents an average of 20 randomly-generated scenarios, and error bars depict the standard deviation of performance over these trials. The bottom figure depicts the average time required to run both algorithms.

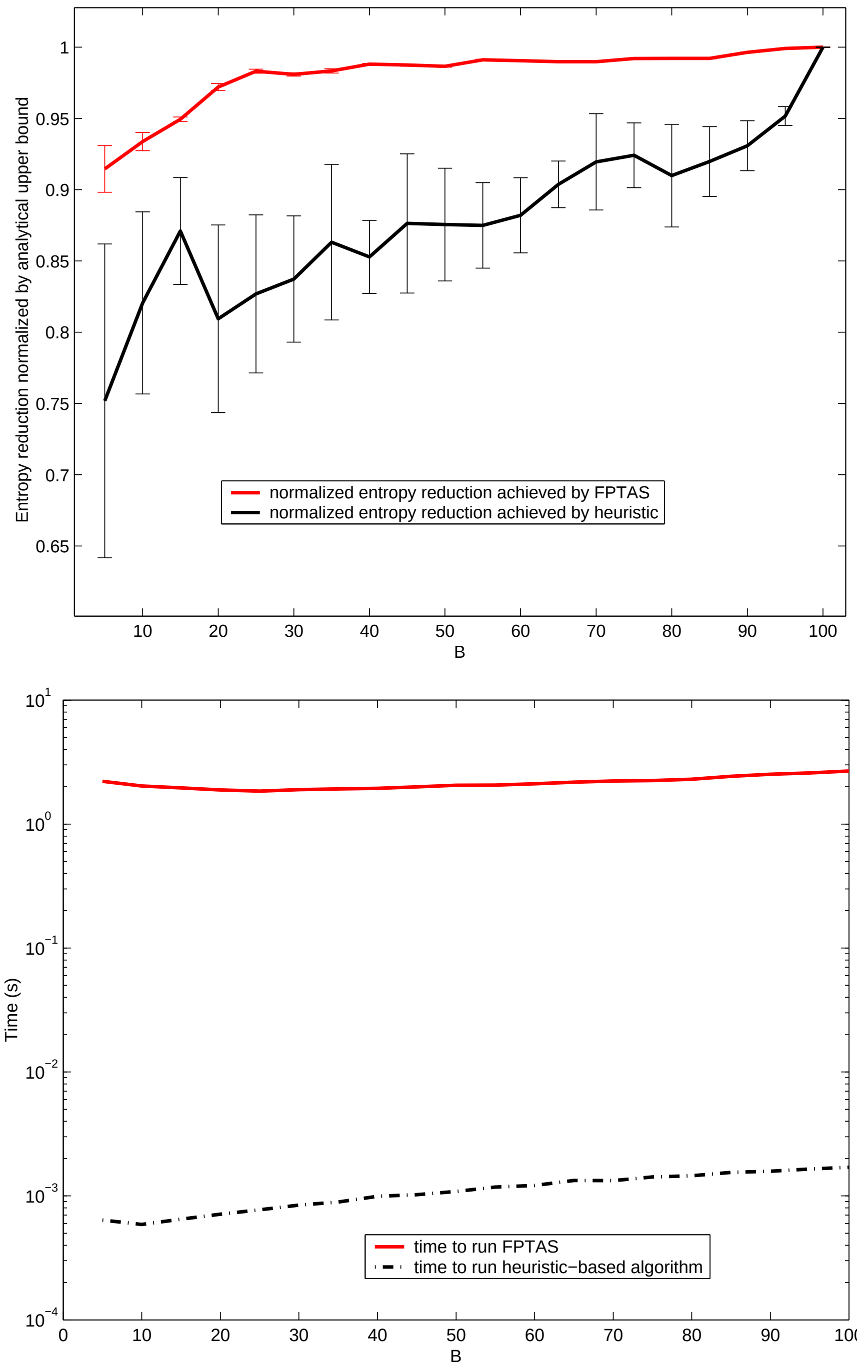

Figure 2. Again the performance of the greedy algorithm is compared to that of the fully polynomial time approximation algorithm, this time for varying B and and a fixed number of features, $m = 20$. The top figure shows the reduction in the entropy of the receiver's estimate achieved by each algorithm, normalized by the maximum possible reduction in entropy as given by the performance bound of the approximation algorithm. Each data point represents an average of 20 randomly-generated scenarios, and error bars depict the standard deviation of performance over these trials. The bottom figure depicts the average time required to run both algorithms.

Two improvements to the approach given in [15] are proposed in this work. First, rather than evaluate the benefit of sending information by what the sender has learned the most about since the previous transmission, this work utilizes an estimate of the receiver's information matrix to evaluate the value of information. It should be noted that such an estimate is not always available, but when available it should be used to the maximum possible benefit. Second, rather than selecting a single large submatrix for transmission, one could instead select multiple smaller submatrices. Allowing this option causes the time complexity of the problem to increase greatly, but it also allows a potentially more intelligent use of scarce communication bandwidth. Thus, even an approximate solution to this more difficult problem may lead to improved performance over the existing simple technique.

Because of the difficulty of solving the more flexible version of the problem, both heuristics and a relaxation approach are utilized in the submatrix-selection algorithm presented in this paper. In order to solve a relaxed version of the problem, work done initially by Vandenberghe *et al.*[25] is leveraged. If the sender's problem as stated in Section IV is relaxed so that the bandwidth constraint and binary decision variables are eliminated, the formulation becomes

$$\begin{aligned} \text{minimize} \quad & c^T x + \log(\det(Y(x)^{-1})) \\ \text{subject to} \quad & Y(x) > 0 \\ & Y(x) = \sum_{i=0}^{M} x_i Y_i \\ & 0 \le x_i \le 1 \qquad i = 0, \ldots, M \\ & \sum_{i=0}^{M} x_i = 1 \end{aligned} \tag{16}$$

This formulation, an extension of the semidefinite programming problem, admits an efficient solution to the relaxed version of the problem under consideration.[7] Note, however, that this formulation includes a decision variable for every possible submatrix of the information matrix. This is an intractably large number for even moderate numbers of map features. Thus, some technique must be found for identifying promising candidate submatrices for consideration in the relaxed algorithm.

Recall that the algorithm in [15] selects a single large submatrix for transmission. We also allow large submatrices to make up some of the transmission candidates. To allow flexibility in the size of candidate submatrices, however, candidates of other sizes are also selected. In this work, large submatrix pairs are also considered in which one submatrix takes up approximately two thirds of the available bandwidth and the other takes up one third, as well as submatrix pairs in which each submatrix takes up approximately half of the available bandwidth.

Selection of these candidate submatrices is accomplished using a stochastic selection heuristic. Candidate submatrices are generated feature by feature in a probabilistic fashion. The first feature is selected according to a probability distribution determined by the magnitudes of the differences between the diagonal elements of the sender's and receiver's information matrices. The next feature is selected according to a distribution based on the magnitude of the cross correlation between the unselected diagonal elements and the previously selected element. Subsequent features are selected in the same fashion, according to a distribution based on the sum of the off-diagonal terms between selected and unselected terms. This process is repeated until the desired number of features for the candidate have been selected. The rationale behind this selection heuristic is that it favors the selection of submatrices containing highly correlated features, which increases the informativeness of the transmitted submatrix.

Transmitting any of these large submatrices (or submatrix pairs) utilizes most of the available bandwidth for cases of interest. However, when significant bandwidth is left over, we use the same stochastic selection heuristic to select multiple smaller candidate submatrices of appropriate size.

Following selection of candidate submatrices, the relaxed problem described above is solved to decide among these candidates. For each set of large submatrices, the relaxed problem is formulated with multiple small submatrices. The small submatrices with the highest weightings in the solution to the relaxed problem are selected for inclusion in a second instance of the relaxed problem, this time subject to the bandwidth constraint. In this second problem, the optimal weightings for the covariance intersection algorithm are found, and a post-transmission entropy is calculated for this set of submatrices. This process is repeated for all large submatrices selected by the stochastic selection heuristic, and the set of large and small submatrices with the lowest post-transmission entropy is selected for transmission.

To summarize, the algorithm proposed for the case of a fully-populated information matrix consists of the following steps:

1. A set of large candidate submatrices is selected using a heuristic method; in this paper, a stochastic metric is considered that is based on the magnitude of differences in the sender's and receiver's information matrices.
2. For each large candidate submatrix (or pair of submatrices), multiple small candidates are generated according to the same heuristic.
3. For each large candidate submatrix (or pair of submatrices) and its associated small submatrices, a relaxed version of the transmission problem is formulated, and this problem is solved with no bandwidth constraint.
4. The small submatrices that received the highest weighting in the relaxed problem are selected to accompany the large submatrix, and the problem is re-solved to find the optimal weights for covariance intersection. Entropy for this set of submatrices is calculated. The submatrices with the lowest post-transmission entropy are ultimately selected for transmission.

Figures 3 and 4 depict the performance of this algorithm. As in Section A, the results of a number of randomly-generated scenarios are shown. In each scenario, sequences of measurements of N features are taken by the sender and the receiver. The sender then uses the algorithm described above to select information for transmission to the receiver. Additionally, a slightly modified version of the algorithm developed in [15] is used as a benchmark. As described above, the features selected in [15] are those about which the sender has accumulated the most information since the previous transmission. Because our formulation considers the results of a single transmission, features in the modified benchmark algorithm are selected based on the magnitude of the difference in information between the sender's and the receiver's information matrices, rather than between the sender's current information matrix and the sender's previous information matrix at the time of the last transmission. This selection process is similar to that of [15] in that a single large submatrix is selected.

The metric by which the algorithms are compared is the reduction in the entropy of the receiver's estimate after incorporating the transmitted submatrices, normalized by the entropy reduction achieved by transmitting submatrices selected through an exhaustive search over all large submatrices. That is, a value of 0.9 in these figures indicates that the algorithm in question resulted in a reduction in the receiver's entropy that was 90% as large as that achieved through exhaustive search; this result would be superior to a value of 0.8, for example.

In Figure 3, the amount of bandwidth available is held constant at $B = 15$ units while the number of map features varies. Note that the algorithm described in this paper achieves reductions in entropy that are very close to those achieved through an exhaustive search, even when no small submatrices are considered. When small submatrices are considered, the algorithm outperforms the exhaustive search over large submatrices by a significant margin. (An exhaustive search including small submatrices was not performed due to the excessive time that would be required.) Although the performance of this algorithm is comparable to that of an exhaustive search, its computation time is dramatically reduced and scales well with problem size.

Figure 4 shows the algorithm's performance for varying availability of communication bandwidth and a fixed number of features ($N = 10$). Again, performance is good, and computation time increases gracefully with problem size.

Figure 5 compares the performance of the algorithm presented in this paper to the performance of the benchmark algorithm based on that in [15] for maps of realistic dimension (N ranging from 10 to 100). These trials are too large to reasonably include an exhaustive search, but the relative performance of the two heuristic algorithms is similar to their performance in the smaller cases, indicating that our algorithm also performs well for large problems. As the figure indicates, the the algorithm developed in this paper consistently achieves up to a 35% greater reduction in entropy than benchmark algorithm.

V. Conclusions and Future Work

The ability to explore and map an unknown environment is seen as a capability of central importance in autonomous robotics, and it is widely acknowledged that in many instances, this task will be done cooperatively and in a communication-limited environment. Yet, the impact of limited communication on the fulfillment of this task has not been adequately studied, nor have algorithms been developed to fully take advantage of the communication bandwidth available.

This work provides insight into the way to best utilize limited communication capabilities in information-sensitive tasks, as well as the performance one can hope to achieve with a limited amount of communication capability. While some work has been done in this area previously, the algorithms presented in this paper provide significant improvements in performance with relatively small increases in computational complexity. In the case of a block-diagonal information matrix, a performance guarantee is given for the algorithm developed in this paper. In the case of a fully-populated information matrix, the algorithm developed in this paper demonstrates good performance in simulation for problems of realistic complexity.

The algorithms described in this paper provide a good solution to the static problem described in Section IV, but they do not address the dynamic decision problem set forth in Section III. Future work will utilize the work described here in the solution of this dynamic decision problem.

Throughout this work it has been assumed that the sender possesses an accurate estimate of the receiver's information matrix, and this has played a key role in the development of the algorithms presented in this paper. In practice, the quality of this estimate is unknown. Future work will examine the degradation of the solution quality with declining accuracy in the sender's estimate. However, in the case of a very poorly-known information matrix, a simple metric such as that of Nettleton *et al.* can still be used in the sequential decision-making process.

Acknowledgments

This work is funded by AFOSR grant FA9550-04-1-0458 and by an NSF Graduate Fellowship.

References

[1]Alighanbari, M. and How, J." Robust Decentralized Task Assignment for Cooperative UAVs." *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2006.

[2]Dash, R.K., Rogers, A., Jennings, N. R., Reece, S. and Roberts, S. "Constrained Bandwidth Allocation in Multi-Sensor Information Fusion: A Mechanism Design Approach", *Proceedings of The Eighth International Conference on Information Fusion*, Philadelphia, PA, 2005.

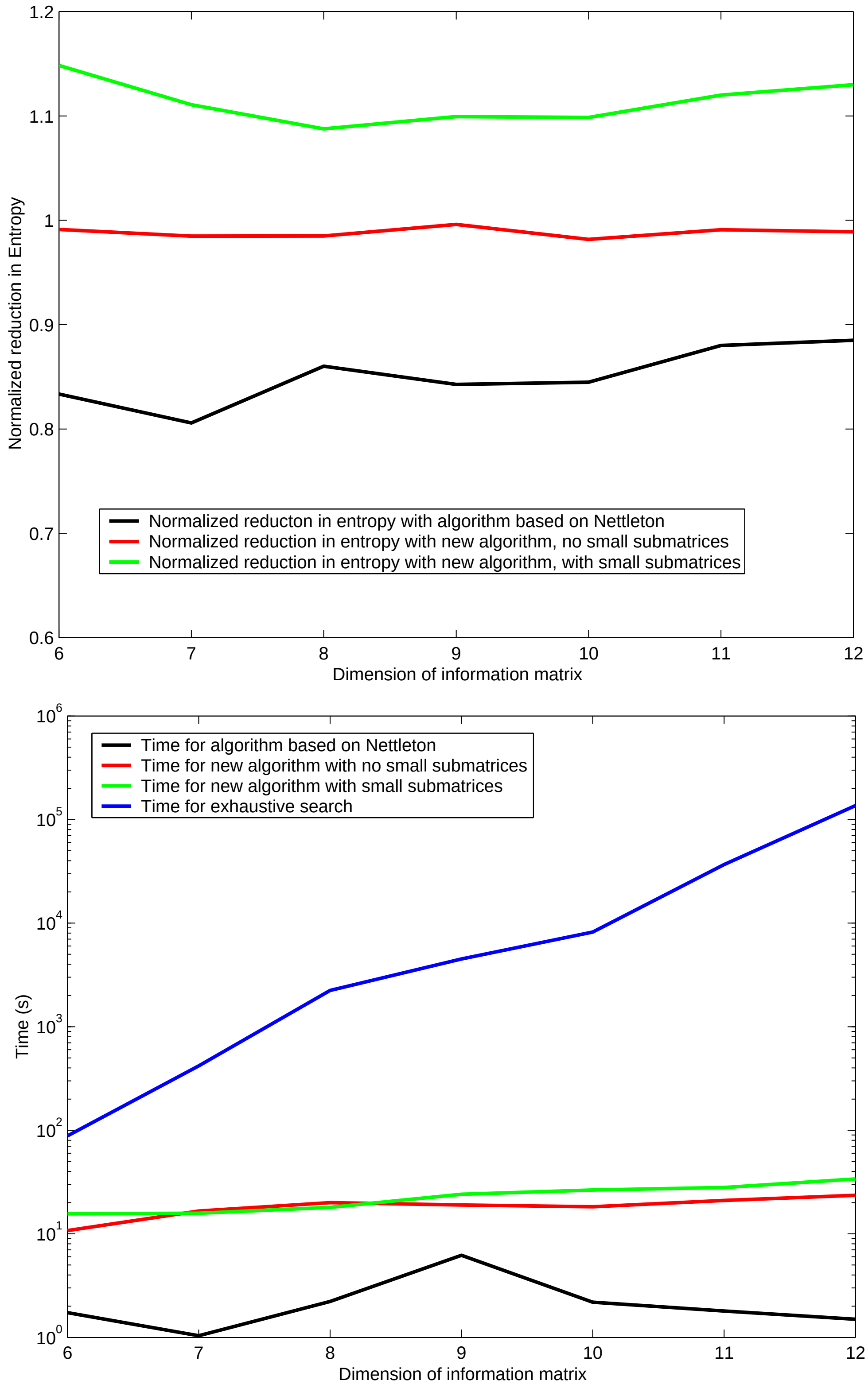

Figure 3. As these simulation results indicate, the algorithm developed in this paper performs very well, reaching entropies very close to those achieved through exhaustive search even when no small submatrices are considered, with a great reduction in computation time. When small submatrices are considered, the algorithm presented in this paper outperforms an exhaustive search over large submatrices with almost no increase in computation time.

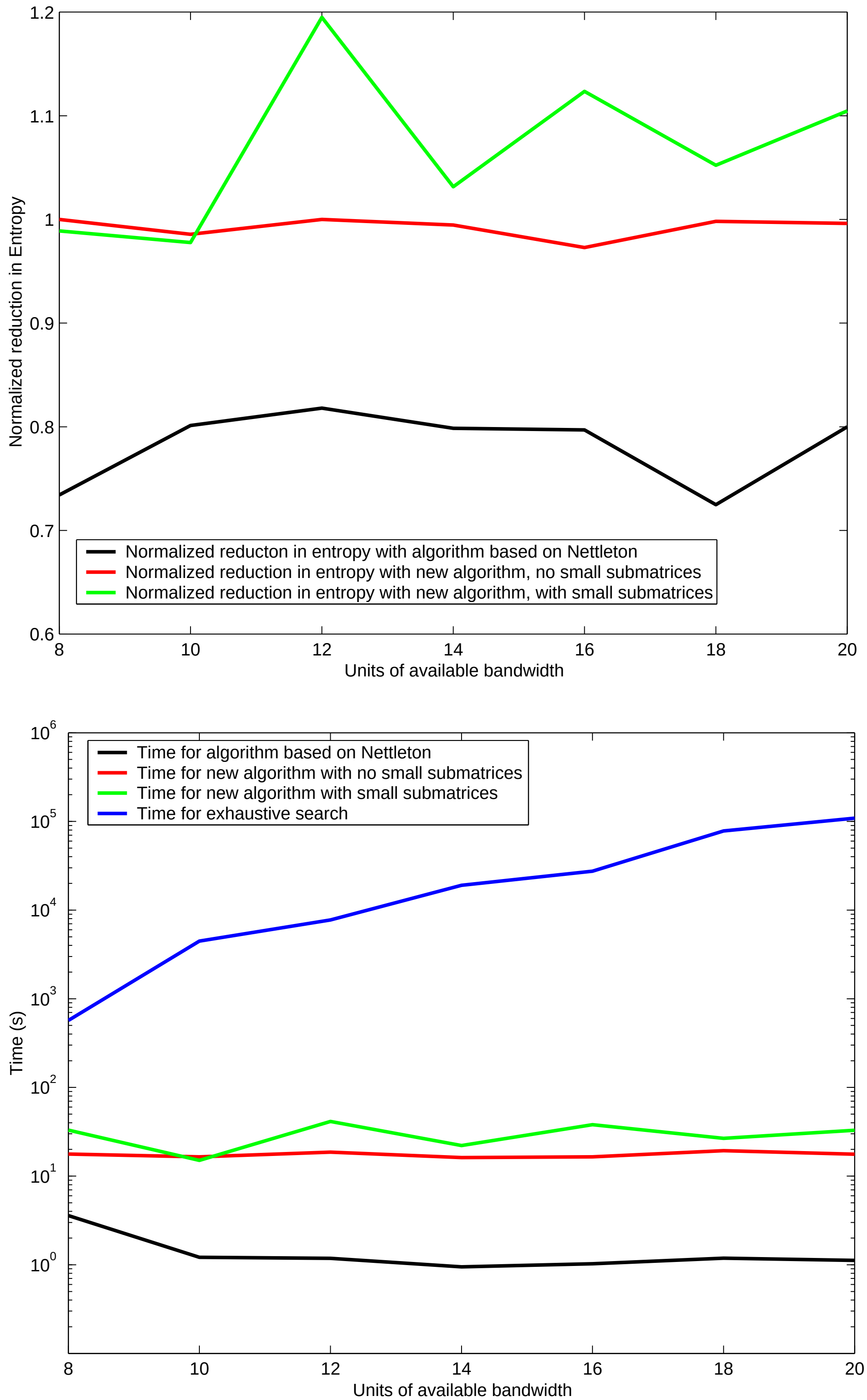

Figure 4. The algorithm developed in Section B also performs well for varying availability of communication bandwidth, and again computation time remains low.

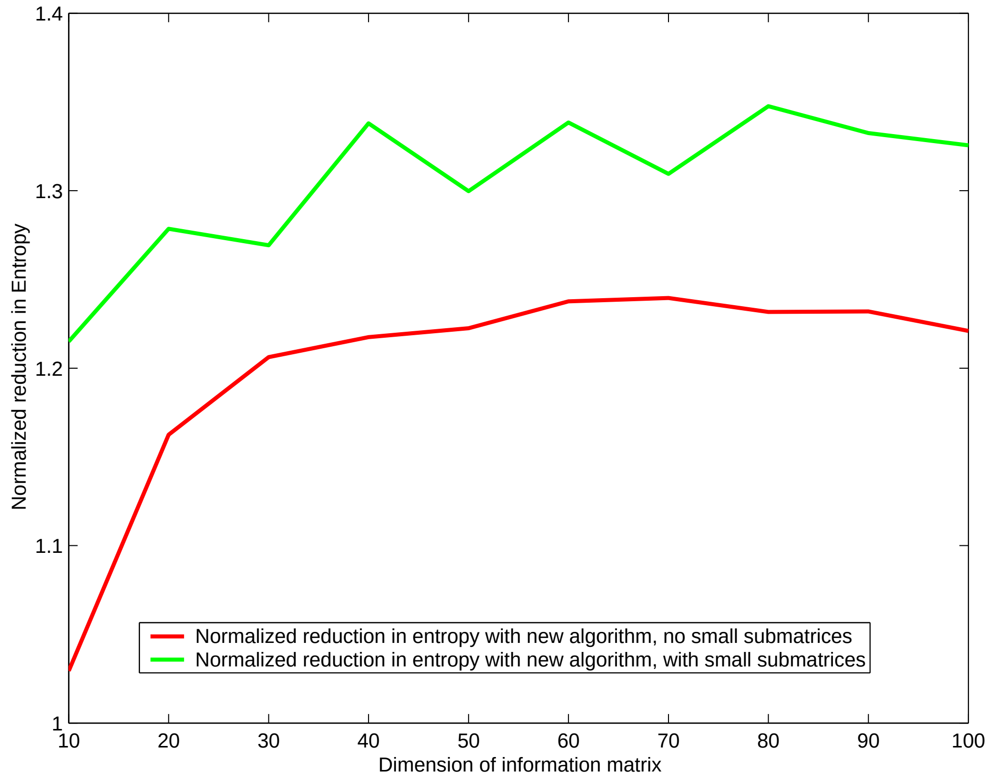

Figure 5. The algorithm presented in this paper continues to outperform the benchmark algorithm as the number of map features and the availability of communication bandwidth increase to realistic magnitudes. Shown is the reduction in entropy obtained by this algorithm, normalized by the entropy reduction obtained by the benchmark algorithm.

[3] Cao, Y. U., Fukunaga, A. S. and Kahng, A. "Cooperative Mobile Robotics: Antecedents and Directions," *Autonomous Robots*, Springer Netherlands, Vol 4, Num. 1, March, 1997.

[4] Cortes, J., Martinez, S., Karatas, T. and Bullo, F. "Coverage control for mobile sensing networks," *IEEE International Conference on Robotics and Automation*, 2002.

[5] Eustice,R., Walter, M. and Leonard, J. "Sparse Extended Information Filters: Insights into Sparsification," In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton, Alberta, Canada, August 2005.

[6] Grocholsky, B. *Information-Theoretic Control of Multiple Sensor Platforms*. PhD thesis, The University of Sydney, 2002.

[7] Julier, S. J. and Uhlmann, J. K. "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," *Proceedings of the American Control Conference*, 1997.

[8] Kellerer, H., Pferschy, U. and Pisinger, D. *Knapsack Problems*. Springer-Verlag, 2004.

[9] Konolige, K., Fox, D., Limketkai, B.,Ko, J. and Stewart, B. "Map merging for distributed robot navigation," *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003.

[10] Kwok, K. S., Driessen, B. J., Phillips, C. A and Tovey, C. A. "Analyzing the Multiple-target-multiple-agent Scenario Using Optimal Assignment Algorithms," *Journal of Intelligent and Robotic Systems* , Vol. 35, No. 1, September, 2002.

[11] Lawler, E. "Fast approximation algorithms for knapsack problems," *Mathematics of Operations Research*, 4:339-356, 1979.

[12] Lawton, J.R.T., Beard, R.W. and Young, B.J. "A decentralized approach to formation maneuvers," *IEEE Transactions on Robotics and Automation*, Vol. 19, Issue: 6. Dec. 2003.

[13] Leonard, J. J. and Durrant-Whyte, H. F. "Simultaneous map building and localization for an autonomous mobile robot." In *Proceedings of the IEEE International Workshop on Intelligent Robots and Systems*, pages 1442-1447, Osaka, Japan, November 1991.

[14]Mutambara, A. G. O. *Decentralized Estimation and Control for Multisensor Systems.* CRC Press, 1998.

[15]Nettleton, E., Thrun, S., Durrant-Whyte, H., and Sukkarieh, H. "Decentralised SLAM with Low-Bandwidth Communication for Teams of Vehicles",*Proceedings of the International Conference on Field and Service Robotics*, Lake Yamanaka, Japan, 2003.

[16]Proakis, J. G. and Salehi, M. *Communication Systems Engineering*, 2nd Edition. Prentice Hall, 2002.

[17]Reece, S. and Roberts, S. "Robust, Low-Bandwidth, Multi-Vehicle Mapping,"*8th International Conference on Information Fusion*, 2005.

[18]Smith, R., Self, M. and Cheeseman, P. "Estimating Uncertain Spatial Relationships in Robotics," *Autonomous Robot Vehicles*, p.167-193, 1990.

[19]Sweeney, J.D., Li, H., Grupen, R. A. and Ramamritham, K. "Scalability and Schedulability in Large, Coordinated, Distributed Robot Systems." *Proceedings of the IEEE International Conference on Robotics and Automation*, Vol. 3, p. 4074- 4079, 2003.

[20]Thrun, S. "Robotic Mapping: A Survey." CMU-CS-02-111, School of Computer Science, Carnegie Mellon University, 2002.

[21]Thrun, S., Burgard, W., and Fox, D. *Probabilistic Robotics*. The MIT Press, Cambridge, MA, 2005.

[22]Thrun, S., Fox, D. and Burgard, W. "A probabilistic approach for concurrent map acquisition and localization for mobile robots." Technical Report CMU-CS-97-183, Carnegie Mellon University, Computer Science Department, Pittsburgh, PA, 1997.

[23]Thrun, S., Liu, Y., Koller, D., Ng, A.Y., Ghahramani, Z. and Durrant-Whyte, H. "Simultaneous localization and mapping with sparse extended information filters." *International Journal of Robotics Research*, 23(7-8), 2004.

[24]Uhlmann, J. K. "General Data Fusion for Estimates with Unknown Cross Covariances," *Proceedings of SPIE, the International Society for Optical Engineering*, 1996.

[25]Vandenberghe, L., Boyd, S. and Wu, S.-P. "Determinant maximization with linear matrix inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, 19(2):499-533, 1998

[26]Walter, M., Eustice, R. and Leonard, J. "A Provably Consistent Method for Imposing Sparsity in Feature-based SLAM Information Filters," *Proceedings of the 12th International Symposium of Robotics Research (ISRR)*, San Francisco, CA, October 2005.

[27]Yamauchi, B. "Frontier-based exploration using multiple robots," in *Proc. of the Second International Conference on Autonomous Agents*, Minneapolis, MN, USA, 1998.

[28]Zlot, R.M., Stentz, A., Dias, M. B. and Thayer, S. "Multi-Robot Exploration Controlled By A Market Economy," *IEEE International Conference on Robotics and Automation*, May, 2002.