Cooperative Sensor Networks with Misinformed Sensors

Zhiyu Yang and Lang Tong
School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853, USA
Email: {zy26, ltong}@ece.cornell.edu

***Abstract*— We consider the communication from a cooperative sensor network to a mobile access point. We assume that sensors are informed with a global message and some nodes are misinformed with random messages. The mobile access point employs a polling-based multiple access control to collect information from the sensor network. We derive the maximum achievable rate for the information retrieval process when d sensors are activated at a time. For the Gaussian multiple access channel under the total network power constraint, we derive an achievable rate expression and show that the maximum achievable rate for the Gaussian multiple access channel is $O(\log_2 d)$, obeying the same scaling law as the capacity of an Gaussian multiple-input-single-output channel.**

I. Introduction

We consider information retrieval in a cooperative SEnsor Network with Mobile Access (SENMA) [1]. As illustrated in Fig. 1, SENMA contains two types of nodes: a large number of low power geographically distributed sensors, and a mobile access point capable of polling sensors individually. By cooperative SENMA (C-SENMA) we mean that, in communicating to the mobile access point, sensors may reach an agreement on the message to transmit, and appropriate codings can be implemented across sensors. This makes the information retrieval robust against failure of individual sensors.

If no sensor is misinformed, and if the mobile access point polls one sensor at a time and the channel between each sensor and the mobile access point is a discrete memoryless channel $q(y|x)$, then the maximum achievable rate for the information retrieval is given by

$$C_0 = \max_{p(x)} I(X;Y).$$

In such a setting, there is no difference between retrieving information from a single sensor or multiple sensors since all sensors have the same message.

For large scale sensor networks, however, reaching complete agreement among all sensors is very difficult, if not impossible. For example, a software agent responsible for distributing the message may not have reached all sensors, or sensors make errors due to unreliable conference links. In practice, there is always a possibility that some sensors do not have the correct message for transmission. We refer to such sensors as misinformed.

The achievable rate of cooperative sensor networks with misinformed sensors is no longer obvious. The error of misinformed sensors cannot be modeled as part of the transmission channel. The reason is that the mobile access point may be able to detect misinformed sensors, and therefore not to poll these nodes again.

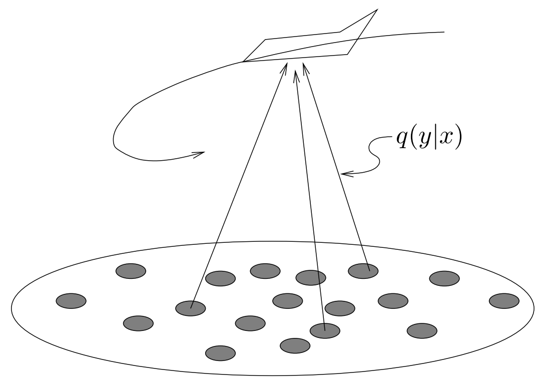


Fig. 1. SEnsor Network with Mobile Access.

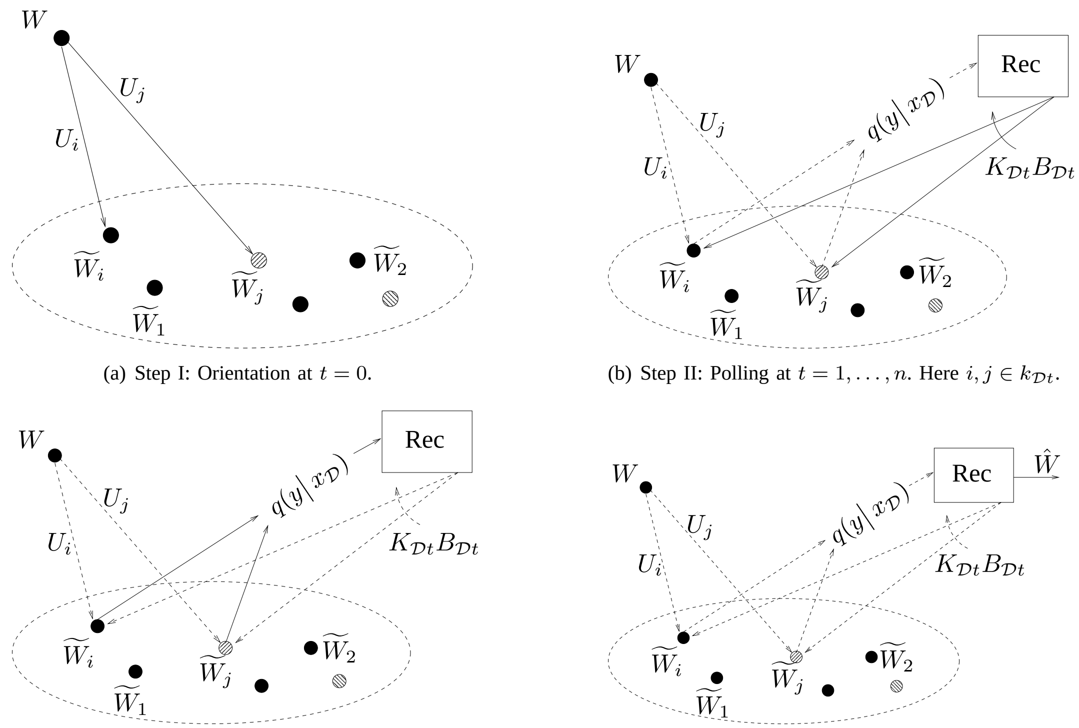
In [2], we investigate a special pulling strategy and characterized the maximum achievable rate. We consider in this paper a general polling strategy that allows the mobile access point to determine which sensors to poll based on previously received transmissions and polling history. It has been proven in [2] that when we activate one sensor at a time, the maximum achievable rate is given by

$$C_1 = (1-\beta)\max_{p(x)} I(X;Y) \tag{1}$$

where β is the probability that the sensor is misinformed. (When the sensor is misinformed, it chooses the message randomly with equal probability.) Notice that if the mobile access point always poll the same sensor, the achievable rate would have been zero. On the other hand, if the mobile access point always polls a new sensor, the achievable rate is lower than the rate in (1).

In this paper, we extend the results to activating d nodes at a time. The maximum achievable rate is given in Theorem 2. We then consider an additive Gaussian multiple access channel with a total power constraint and show that the maximum achievable rate is $O(\log_2 d)$, obeying the same rule as the case when there is no misinformed sensors.

For notation compactness, denote an entry with two sub-indices i and j by $(\cdot)_{ij}$. The meaning of ij, a double-index or a scalar multiplication, can be determined by its context.



(a) Step I: Orientation at $t = 0$.

(b) Step II: Polling at $t = 1, \dots, n$. Here $i, j \in k_{\mathcal{D}t}$.

(c) Step III: Transmission and Reception at $t = 1, \dots, n$.

(d) Step IV: Decoding at $t = n$.

Fig. 2. Communication steps.

II. Model

The communication of the global message from the network to the mobile access point is divided into four steps as shown in Fig. 2: (a) orientation, (b) polling, (c) transmission and reception, and (d) decoding. In the first step, nodes are informed with the globe message $W \in \{1, \dots, M\}$ that is uniformly distributed. Due to the size of the network, a node may be informed incorrectly and end up with a different message. We assume that each node receives the globe message correctly with some probability, and the reception is independent of other nodes. More specifically, the reception of node i is controlled by a binary random variable U_i, independent of W and identically independently distributed (i.i.d.) across node index i with distribution

$$p(u_i) = \begin{cases} \beta & \text{if } u_i = 0 \\ 1-\beta & \text{if } u_i = 1 \end{cases}$$

where $\beta \in [0, 1]$ is a constant. When $U_i = 1$, the received message at node i, $\widetilde{W}_i$, is equal to the globel message W. When $U_i = 0$, $\widetilde{W}_i$ is uniformly distributed from 1 to M. Thus

$$p(\tilde{w}_i \big| w, u_i) = \begin{cases} \delta(\tilde{w}_i, w) & \text{if } u_i = 1 \\ \frac{1}{M} 1_{1 \le \tilde{w}_i \le M} & \text{if } u_i = 0 \end{cases}$$

where $\delta(a, b)$ is equal to 1 if $a = b$, 0 otherwise, and the indicating function 1_A equal to 1 if event A is true, 0 otherwise. The constant β controls the reception of the globe message by individual nodes and is referred to as the orientation error probability of the network.

The mobile access point comes to retrieve information from the field after the information orientation has been accomplished. The information retrieval process consists of Step 2 Polling and Step 3 Transmission and Reception. We assume the mobile access point has the ability to poll individual sensors and a polling-based multiple access is employed: at each time slot, the mobile access point polls d nodes, each transmitting one symbol to the channel. Let $\mathcal{D} \triangleq \{1, \dots, d\}$ and $Z_{\mathcal{D}}$ denote $\{Z_1, \dots, Z_d\}$ where $Z_1, \dots, Z_d$ are generic symbols. We assume each node has its own code book. At time t, the access point polls nodes $K_{\mathcal{D}t} = \{K_{1t}, \dots, K_{dt}\}$ and symbols $B_{\mathcal{D}t} = \{B_{1t}, \dots, B_{dt}\}$. More specifically, for $j \in \mathcal{D}$, node K_{jt} transmits the B_{jt}-th symbol of the codeword corresponding to the $\widetilde{W}_{K_{jt}}$-th message in node K_{jt}'s code book. Since the mobile access point is usually equipped with high-gain antennas and high-power transmitter, we assume the polling channel is error free.

We assume that the uplink multiple access channels (MACs) from any d nodes to the mobile access point are identical and modeled by a discrete memoryless MAC $\{\mathcal{X}, \mathcal{Y}, q(y \big| x_{\mathcal{D}})\}$, where $\mathcal{X}$ and $\mathcal{Y}$ are the input and output alphabets respectively, and $q(y \big| x_{\mathcal{D}})$ is the transition probability of the channel. We assume the MAC is symmetrical with respect to input permutations, *i.e.*, $q(y \big| x_{\pi_1}, \dots, x_{\pi_d})$ is identical for all permutations $\pi_{\mathcal{D}}$ in the domain $\mathcal{D}$.

For $j \in \mathcal{D}$ and at time t, node K_{jt}, after receiving the polling signal, transmits the selected symbol to the uplink channel. Denote X_{jt} the transmission from node K_{jt} and Y_t the output of the MAC at time t. After receiving Y_t, the mobile access point moves to the next time slot $t+1$ and starts the polling step again. It may poll nodes that have or have not been polled before. Step 2 and Step 3 alternate until t reaches n, the number of time slots the mobile access point spends to retrieve information from the field.

After polling some nodes, the access point may have gained some knowledge about whether the nodes polled have received the globe message correctly in the orientation step. If it believes that one node is misinformed, it is then beneficial not to poll that node again. We therefore allow the polling signal $K_{\mathcal{D}t}, B_{\mathcal{D}t}$ at time t to depend on the previous polling signals $K_{\mathcal{D}}^{t-1}, B_{\mathcal{D}}^{t-1}$ and the previous received channel outputs Y^{t-1}.

In the last step, the access point decodes the globe message based on the observation of Y^n and the polling history $K_{\mathcal{D}}^n, B_{\mathcal{D}}^n$. The decoded message is denoted by $\hat{W} \in \{1, \ldots, M\}$. Let $W_{jt} \triangleq \widetilde{W}_{K_{jt}}$ be the message at the jt-th polled node. The communication process described above is summarized in Fig. 3.

We assume that the sensor network is large in the sense that there are infinite number of nodes, and there is no limit on how many times a node can be polled.

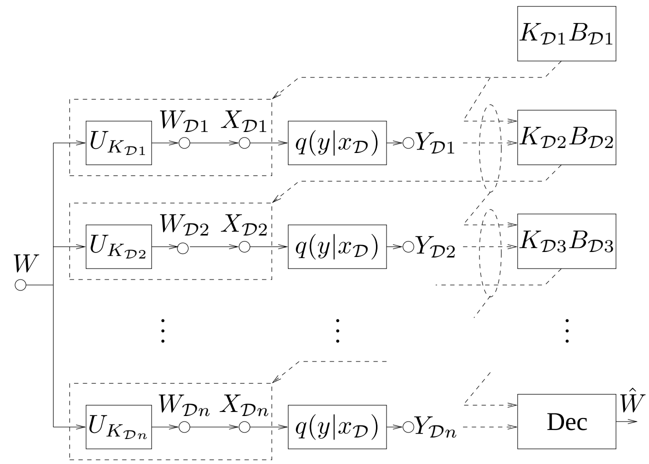


Fig. 3. Channel model.

III. Maximum Achievable Rate with Polling-based Multiple Access: Definitions and Results

Denote $(\beta, \mathcal{X}, \mathcal{Y}, q(y|\, x_{\mathcal{D}}))$ the cooperative SENMA (C-SENMA) with infinite nodes, orientation error probability β, and communication MAC $(\mathcal{X}, \mathcal{Y}, q(y|\, x_{\mathcal{D}}))$. Let $\mathcal{N}$ be the set of positive numbers and $\mathcal{W} \triangleq \{1, 2, \ldots, M\}$. A code book with M messages, denoted by $\mathbb{C}$ or $c(k, b, w)$, is a mapping from $\mathcal{N} \times \mathcal{N} \times \mathcal{W}$ to $\mathcal{X}$, where the first $\mathcal{N}$ (or k) represents the node index, the second $\mathcal{N}$ (or b) the symbol index, and $\mathcal{W}$ the message index. That is, $c(k, b, w)$ represents the b-th symbol of the w-th codeword at the k-th node.

A polling policy, or policy in short, is a set of conditional distributions $\mathbb{P} = \{q_t(k_{\mathcal{D}t}, b_{\mathcal{D}t}|\, k_{\mathcal{D}}^{t-1}, b_{\mathcal{D}}^{t-1}, y^{t-1}) : t \geq 1\}$, where $k_{jt}, b_{jt} \geq 1$ and $y_t \in \mathcal{Y}$ for all $t \geq 1$ and all $j \in \mathcal{D}$. After polling $(k_{\mathcal{D}}^{t-1}, b_{\mathcal{D}}^{t-1})$ and receiving y^{t-1}, the mobile access point generates the polling signal $(K_{\mathcal{D}t}, B_{\mathcal{D}t})$ randomly according to the conditional distribution $q_t(k_{\mathcal{D}t}, b_{\mathcal{D}t}|\, k_{\mathcal{D}}^{t-1}, b_{\mathcal{D}}^{t-1}, y^{t-1})$. Due to the symmetry of the sensor network, without loss of generality, we assume that $q_t(k_{jt}|\, k_{\mathcal{D}}^{t-1}, b_{\mathcal{D}}^{t-1}, y^{t-1}) = 0$ if $k_{jt} > dt$ for $j \in \mathcal{D}$, *i.e.*, we only consider sensors 1 to dt at time t.

A decoder with n channel uses and M messages, denoted by $\mathbb{D}$ or $d(k_{\mathcal{D}}^n, b_{\mathcal{D}}^n, y^n)$, is a mapping from $\mathcal{N}^{dn} \times \mathcal{N}^{dn} \times \mathcal{Y}^n$ to $\mathcal{W}$. The mobile access point decodes the globe message by assigning $\hat{w} = d(k_{\mathcal{D}}^n, b_{\mathcal{D}}^n, y^n)$, where $(k_{\mathcal{D}}^n, b_{\mathcal{D}}^n)$ is the polling history and y^n is the received channel outputs.

For a given triple $(\mathbb{C}, \mathbb{P}, \mathbb{D})$, the rate of communication is defined as $R \triangleq \log(M)/n$, where M is the number of messages of the code book and n is the number of channel uses of the decoder.

For a given C-SENMA, the error probability of a communication triple $(\mathbb{C}, \mathbb{P}, \mathbb{D})$ is defined as $P_e \triangleq \mathcal{P}(\hat{W} \neq W)$, where $W \in \{1, \cdots, M\}$ is uniformly distributed and $\hat{W}$ is the decoded message with the communication triple.

For a C-SENMA, a rate R is achievable if there exists a sequence of $(\mathbb{C}, \mathbb{P}, \mathbb{D})$ triples with rate R such that the corresponding error probabilities go to zero.

As a special case when $d = 1$, the maximum achievable rate is given by the following theorem:

Theorem 1 ([2]): For $d = 1$, the maximum achievable rate of a C-SENMA $(\beta, \mathcal{X}, \mathcal{Y}, q(y|\, x))$ with polling-based multiple access is

$$C_1 = (1-\beta) \max_{p(x)} I(X; Y),$$

where $X \in \mathcal{X}$, $Y \in \mathcal{Y}$, and $p(y|\, x) = q(y|\, x)$.

The extension of Theorem 1 is summarized as follows:

Theorem 2: For a given d, consider random variables $X_{\mathcal{D}}$ with distribution $p(x_{\mathcal{D}})$. Denote $p_{\mathcal{I}}(x_{\mathcal{I}})$ the marginal distribution of $X_{\mathcal{I}}$ for $\mathcal{I} \subset \mathcal{D}$. The maximum achievable rate of a C-SENMA $(\beta, \mathcal{X}, \mathcal{Y}, q(y|\, x_{\mathcal{D}}))$ with polling-based multiple access is

$$C_d = \max_{p(x_{\mathcal{D}})} \sum_{\mathcal{I} \subset \mathcal{D}} (1-\beta)^{|\mathcal{I}|} \beta^{d-|\mathcal{I}|} I(X_{\mathcal{I}}^{(\mathcal{I})}; Y),$$

where $X_{\mathcal{D}}^{(\mathcal{I})} = (X_1^{(\mathcal{I})}, \ldots, X_d^{(\mathcal{I})})$ are derived from $X_{\mathcal{D}}$ with distribution

$$p^{(\mathcal{I})}(x_{\mathcal{D}}^{(\mathcal{I})}) = p_{\mathcal{I}}(x_{\mathcal{I}}^{(\mathcal{I})}) \prod_{j \notin \mathcal{I}} p_j(x_j^{(\mathcal{I})}),$$

and $p(y|\, x_{\mathcal{D}}^{(\mathcal{I})}) = q(y|\, x_{\mathcal{D}}^{(\mathcal{I})})$.

The direct part of Theorem 2 is outlined in the appendix. Due to the space limit, the converse part is omitted here.

IV. Gaussian Multiple Access Channels with a Total Power Constraint

Consider the following Gaussian multiple access channel

$$y = v + \sum_i x_i \tag{2}$$

where $x_i \in \mathcal{C}$ is the input from the i-th sensor, $v \in \mathcal{C}$ is the additive white Gaussian noise with zero mean and unit variance, and $y \in \mathcal{C}$ is the channel output. We impose a total power constraint P on the network, *i.e.*, the total transmitted power from all sensors is less than or equal to P:

$$\frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{d}|x_{jt}|^2 \le P$$

where x_{jt} is the jt-th transmission. If there is no orientation error, *i.e.*, $\beta = 0$, we know from the multiple-input-single-output (MISO) channel capacity that, with d sensor polled at a time, the maximum achievable rate is $C_d^{(0)} = \log_2(1+dP)$. Therefore, when $\beta = 0$, the maximum achievable rate $C_d^{(0)} = O(\log_2 d)$ goes to infinity as we increase d, the number of sensors polled at a time. In this section, we show that even with $\beta > 0$, the maximum achievable rate of the channel (2) is still $O(\log_2 d)$.

Consider activating d sensors at a time. Let the input random variables $X_1, \ldots, X_d$ be identically Gaussian distributed $\mathcal{N}_C(0, P/d)$ and let any two input random variables X_i, X_j have correlation coefficient 1. For $\mathcal{I} \subset \mathcal{D}$, since the derived random variables $X_{\mathcal{D}\setminus\mathcal{I}}^{(\mathcal{I})}$, independent of each other, are independent of $X_{\mathcal{I}}^{(\mathcal{I})}$, $X_{\mathcal{D}\setminus\mathcal{I}}^{(\mathcal{I})}$ contribute $(d-|\mathcal{I}|)P/d$ power to the additive noise. Therefore,

$$I(X_{\mathcal{I}}^{(\mathcal{I})};Y) = \log_2\Big(1 + \frac{|\mathcal{I}|^2 P/d}{1+(d-|\mathcal{I}|)P/d}\Big).$$

From Theorem 2, the following rate is achievable,

$$R_d = \sum_{i=0}^{d}(1-\beta)^i\beta^{d-i}\binom{d}{i}\log_2\Big(1+\frac{i^2P/d}{1+(d-i)P/d}\Big).$$

The next proposition shows that $R_d = O(\log_2 d)$. Since $R_d \le C_d \le C_d^{(0)} = O(\log_2 d)$, we have $C_d = O(\log_2 d)$.

Proposition 3: For $\beta \in [0,1)$ and $P > 0$,

$$\lim_{d\to\infty}(R_d - \log_2 d) = \log_2\Big(\frac{(1-\beta)^2P}{1+\beta P}\Big).$$

Proof: Let $S_1, S_2, \ldots,$ be i.i.d. Bernoulli with mean $1-\beta$. Then $T_d = \sum_{i=1}^{d} S_i$ is binomial distributed. Let

$$f(a,b) \triangleq \log_2\Big(a + \frac{b^2P}{1+(1-b)P}\Big).$$

We have

$$\begin{aligned}R_d - \log_2 d &= \sum_{i=0}^{d}(1-\beta)^i\beta^{d-i}\binom{d}{i} f(1/d, i/d)\\ &= E[f(1/d, T_d/d)].\end{aligned}$$

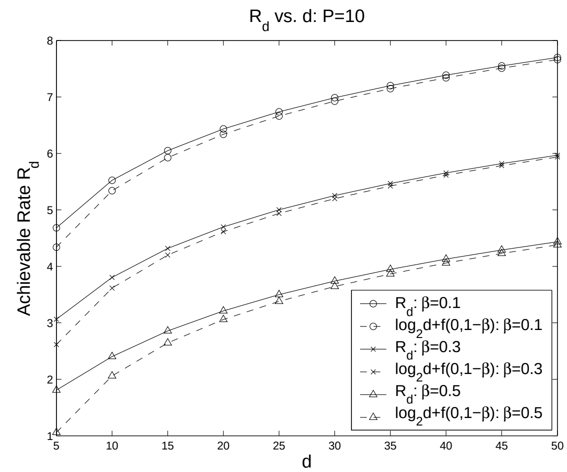


Fig. 4. Achievable rate R_d versus d: $P = 10$ and $\beta = 0.1, 0.3, 0.5$.

Since $f(a,b)$ is continues at $(0, 1-\beta)$, for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $a \in [0,\delta) \triangleq \mathcal{A}$ and $b \in (1-\beta-\delta, 1-\beta+\delta) \triangleq \mathcal{B}$, $|f(a,b) - f(0,1-\beta)| \le \epsilon/3$. For the same ϵ and δ,

$$\begin{aligned}&|E[f(1/d, Y_d/d)] - f(0,1-\beta)|\\ &\le E[|f(1/d,Y_d/d) - f(0,1-\beta)|]\\ &\le E_{Y_d/d\in\mathcal{B}}[|f(1/d,Y_d/d) - f(0,1-\beta)|]\\ &\quad + E_{Y_d/d\notin\mathcal{B}}[|f(1/d,Y_d/d)|]\\ &\quad + \mathcal{P}_r(Y_d/d\notin\mathcal{B}) f(0,1-\beta).\end{aligned} \tag{3}$$

For large d, $1/d \in \mathcal{A}$. Therefore, due to the continuity of f, the first term in (3) is upper bounded by $\epsilon/3$. The third term, by the weak law of large number, is upper bounded by $\epsilon/3$ for large d. We bound the second term as follows. For $d \ge 1$ and $i = 0, \ldots, d$, we have

$$\begin{aligned}f(1/d, i/d) &\le \log_2(1/d + P),\\ f(1/d, i/d) &\ge \log_2(1/d).\end{aligned}$$

For large d such that $1/d + P \le d$, we have

$$|f(1/d, i/d)| \le \log_2 d.$$

Hence, by Chebyshev inequality,

$$\begin{aligned}E_{Y_d/d\notin\mathcal{B}}[|f(1/d,Y_d/d)|] &\le \mathcal{P}_r(Y_d/d\notin\mathcal{B})\log_2 d\\ &\le \frac{\sigma_s^2}{\delta^2 d}\log_2 d\\ &\le \epsilon/3 \qquad \text{for large } d.\end{aligned}$$

Therefore, $|E[f(1/d,Y_d/d)] - f(0,1-\beta)| \le \epsilon$ for large d. Since $\epsilon > 0$ is arbitrary, $\lim_{d\to\infty} E[f(1/d, Y_d/d)] = f(0, 1-\beta)$, proving the proposition. □

Fig. 4 shows the achievable rate R_d and the approximation function $\log_2 d + f(0, 1-\beta)$ versus d when $P = 10$ and $\beta = 0.1, 0.3, 0.5$. As shown in Fig. 4, R_d and the approximation function converge as d increases. As expected, the achievable rate is higher for a network with a smaller orientation error probability.

V. SUMMARY

We have presented the maximum achievable rate for cooperative sensor networks with misinformed sensors when d sensors are activated at a time. We have considered an additive Gaussian multiple access channel with a network power constraint and shown that the maximum achievable rate is $O(\log_2 d)$.

APPENDIX
SKETCH OF THE DIRECT PART

The communication of the C-SENMA involves the design of the $(\mathbb{C}, \mathbb{P}, \mathbb{D})$ triple. We first derive an achievable rate based on a repetitive polling policy, the N-polling policy, and then optimize the achievable rate to proof the achievability of Theorem 2.

The N-polling policy is a deterministic policy that groups every N time slots into a time frame (Fig. 5) and polls in a time frame d nodes that have never been polled before, regardless of the received channel outputs. More specifically, we active nodes $(i-1)d+1, \ldots, id$ in the i-th time frame. And for node $(i-1)d+j$, $j = 1, \ldots, d$, we poll the $((t-1)d+j)$-th symbol at time $t = 1, \ldots, n$. Mathematically, in the N-polling policy,

$$
\begin{aligned}
&q_t(k_{\mathcal{D}t}, b_{\mathcal{D}t} \big|\, k_{\mathcal{D}}^{t-1}, b_{\mathcal{D}}^{t-1}, y^{t-1}) \\
&= \begin{cases} 1 & \text{if } k_{jt} = \lfloor (t-1)/N \rfloor d + j \text{ and } b_{jt} = (t-1)d + j \\ & \text{for all } j = 1, \ldots, d \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

where $\lfloor a \rfloor$ is the largest integer not greater than a.

Lemma 4: For a given d, consider random variables $S_{\mathcal{D}} \in \mathcal{X}^d$ with distribution $p(s_{\mathcal{D}})$. Denote $p_{\mathcal{I}}(s_{\mathcal{I}})$ the marginal distribution of $S_{\mathcal{I}}$ for $\mathcal{I} \subset \mathcal{D}$. Let $\bar{\mathcal{I}} \triangleq \mathcal{D} \setminus \mathcal{I}$. For a given C-SENMA $(\beta, \mathcal{X}, \mathcal{Y}, q(y \big|\, x_{\mathcal{D}}))$, with the 1-polling policy, the rate

$$R_{d1} \triangleq I(S_{\mathcal{D}}; Y)$$

is achievable, where

$$p(y \big|\, s_{\mathcal{D}}) = \sum_{\mathcal{I} \subset \mathcal{D}} (1-\beta)^{|\mathcal{I}|} \beta^{|\bar{\mathcal{I}}|} \sum_{s'_{\bar{\mathcal{I}}} \in \mathcal{X}^{|\bar{\mathcal{I}}|}} q(y \big|\, s_{\mathcal{I}}, s'_{\bar{\mathcal{I}}}) \prod_{j \in \bar{\mathcal{I}}} p_j(s'_j).$$

Proof: Code book Generation: Generate a code book with $M = 2^{nR}$ messages at random according to the distribution $p(s_{\mathcal{D}})$. Specifically, for $1 \leq t \leq n$, let $s_{\mathcal{D}t}(w) = (s_{1t}(w), \ldots, s_{dt}(w))$ be a mapping from $\mathcal{W}$ to $\mathcal{X}^d$. For $1 \leq t \leq n$, $1 \leq w \leq M$, assign $s_{\mathcal{D}t}(w)$ a value independently generated according to the distribution $p(s_{\mathcal{D}})$. After the random assignment of $s_{\mathcal{D}t}(w)$, let every node have the same code book, *i.e.*, let $c(k, (t-1)d+j, w) = s_{jt}(w)$ for all $1 \leq k \leq nd$, $1 \leq t \leq n$, $1 \leq j \leq d$, and $1 \leq w \leq M$.

Decoder: Typical set decoding is employed. Define the typical set $A_\epsilon^{(n)}$ with respect to the distribution $p(s_{\mathcal{D}}, y)$

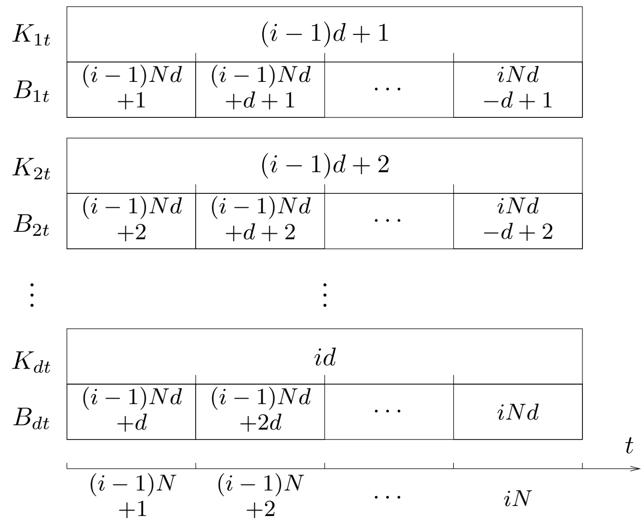


Fig. 5. N-polling signals in the i-th time frame.

$$
\begin{aligned}
A_\epsilon^{(n)} \triangleq \{ (s_{\mathcal{D}}^n, y^n) \in \mathcal{X}^{dn} \times \mathcal{Y}^n :& \\
\Big| -\frac{1}{n} \log p(s_{\mathcal{D}}^n) - H(S_{\mathcal{D}}) \Big| &\leq \epsilon, \\
\Big| -\frac{1}{n} \log p(y^n) - H(Y) \Big| &\leq \epsilon, \\
\Big| -\frac{1}{n} \log p(s_{\mathcal{D}}^n, y^n) - H(S_{\mathcal{D}}, Y) \Big| &\leq \epsilon \},
\end{aligned}
$$

where $p(s_{\mathcal{D}}^n, y^n) = \prod_{i=1}^n p(s_{\mathcal{D}i}, y_i)$. Upon receiving channel outputs y^n, the mobile access point declares the message $\hat{w}$ as the received message if there is one and only one $\hat{w} \in \mathcal{W}$ such that $(s_{\mathcal{D}}^n(\hat{w}), y^n) \in A_\epsilon^{(n)}$; otherwise, the decoder declares an error.

The error analysis is omitted here due to the space limit. It can be shown that, if $R < I(S_{\mathcal{D}}; Y)$, the average probability of error, average over all codewords and all code books, converges to zero as n goes to infinity. $\square$

Corollary 5: For a given d, a distribution $p(s_{\mathcal{D}})$ on the alphabet $\mathcal{X}^d$, and a C-SENMA $(\beta, \mathcal{X}, \mathcal{Y}, q(y \big|\, x_{\mathcal{D}}))$, with the N-polling policy, the rate

$$R_{dN} \triangleq \frac{1}{N} I(S_{\mathcal{D}}^N; Y^N)$$

is achievable, where $S_{\mathcal{D}}^N \in \mathcal{X}^{dN}$, $Y^N \in \mathcal{Y}^N$, and

$$
\begin{aligned}
p(s_{\mathcal{D}}^N, y^N) = \Big(\prod_{i=1}^N p(s_{\mathcal{D}i}) \Big) \Big(\sum_{\mathcal{I} \subset \mathcal{D}} (1-\beta)^{|\mathcal{I}|} \beta^{|\bar{\mathcal{I}}|} \\
\cdot \prod_{i=1}^N \sum_{s'_{\bar{\mathcal{I}}} \in \mathcal{X}^{|\bar{\mathcal{I}}|}} q(y_i \big|\, s_{\mathcal{I}i}, s'_{\bar{\mathcal{I}}}) \prod_{j \in \bar{\mathcal{I}}} p_j(s'_j) \Big). \quad (4)
\end{aligned}
$$

Proof: Consider the N-th extension of the MAC, $(\mathcal{X}^N, \mathcal{Y}^N, q(y^N \big|\, x_{\mathcal{D}}^N))$, where

$$q(y^N \big|\, x_{\mathcal{D}}^N) = \prod_{i=1}^N q(y_i \big|\, x_{\mathcal{D}i}).$$

Let $p(s_{\mathcal{D}}^N)=\prod_{i=1}^N p(s_{\mathcal{D}i})$ be the input distribution to the N-th extended C-SENMA $(\beta,\mathcal{X}^N,\mathcal{Y}^N,q(y^N|\, x_{\mathcal{D}}^N))$. By Lemma 4, $I(S_{\mathcal{D}}^N;Y^N)$ is achievable for the N-th extended system with 1-polling, where

$$\begin{aligned} p(s_{\mathcal{D}}^N,y^N) &= \Big(\prod_{i=1}^N p(s_{\mathcal{D}i})\Big)\Big(\sum_{\mathcal{I}\subset\mathcal{D}}(1-\beta)^{|\mathcal{I}|}\beta^{|\bar{\mathcal{I}}|} \\ &\quad\cdot\sum_{s_{\bar{\mathcal{I}}}'^N\in\mathcal{X}^{N|\bar{\mathcal{I}}|}}\prod_{i=1}^N q(y_i|\, s_{\mathcal{I}i},s_{\bar{\mathcal{I}}i}')\prod_{j\in\bar{\mathcal{I}}}p_j(s_{ji}')\Big) \\ &= \Big(\prod_{i=1}^N p(s_{\mathcal{D}i})\Big)\Big(\sum_{\mathcal{I}\subset\mathcal{D}}(1-\beta)^{|\mathcal{I}|}\beta^{|\bar{\mathcal{I}}|} \\ &\quad\cdot\prod_{i=1}^N\sum_{s_{\bar{\mathcal{I}}}'\in\mathcal{X}^{|\bar{\mathcal{I}}|}}q(y_i|\, s_{\mathcal{I}i},s_{\bar{\mathcal{I}}}')\prod_{j\in\bar{\mathcal{I}}}p_j(s_j')\Big). \end{aligned}$$

The operation of the original C-SENMA with the N-polling policy is equivalent to that of the N-th extended C-SENMA with the 1-polling policy. Thus $\frac{1}{N}I(S_{\mathcal{D}}^N;Y^N)$ is achievable for the original system with N-polling. □

We are now ready to prove the achievability of Theorem 2 by showing the convergence of R_{dN} as N goes to infinity. Let $S_{\mathcal{D}}^N$ has distribution $\prod_{i=1}^N p(s_{\mathcal{D}i})$. Introduce $V_{\mathcal{D}}\in\{0,1\}^d$, i.i.d. Bernoulli random variables with mean $1-\beta$. Let $V_{\mathcal{D}}$ be independent of $S_{\mathcal{D}}^N$. It can be shown that if we let

$$p(y^N|\, s_{\mathcal{D}}^N,v_{\mathcal{D}})=\prod_{i=1}^N\sum_{s_{\bar{\mathcal{I}}}'\in\mathcal{X}^{|\bar{\mathcal{I}}|}}q(y_i|\, s_{\mathcal{I}i},s_{\bar{\mathcal{I}}}')\prod_{j\in\bar{\mathcal{I}}}p_j(s_j') \quad (5)$$

where $\mathcal{I}=\{i:v_i=1\}$, then the resulting marginal distribution $p(s_{\mathcal{D}}^N,y^N)$ is given by (4).

Therefore,

$$\begin{aligned} \frac{1}{N}I(S_{\mathcal{D}}^N;Y^N) &= \frac{1}{N}I(S_{\mathcal{D}}^N;Y^N,V_{\mathcal{D}})-\frac{1}{N}I(S_{\mathcal{D}}^N;V_{\mathcal{D}}|\, Y^N) \\ &\geq \frac{1}{N}I(S_{\mathcal{D}}^N;Y^N|\, V_{\mathcal{D}})-\frac{d}{N} \qquad (6) \\ &\to \frac{1}{N}I(S_{\mathcal{D}}^N;Y^N|\, V_{\mathcal{D}}) \qquad \text{as } N\to\infty \end{aligned}$$

where (6) is because $H(V_{\mathcal{D}})\leq d$. For all $v_{\mathcal{D}}\in\{0,1\}^d$, let $\mathcal{I}=\{i:v_i=1\}$. Let $X_{\mathcal{D}}^{(\mathcal{I})}$, derived from $p(s_{\mathcal{D}})$, and $\mathcal{Y}$ be defined as in Theorem 2. It can be shown from (5) that

$$\frac{1}{N}I(S_{\mathcal{D}}^N;Y^N|\, v_{\mathcal{D}})=I(X_{\mathcal{I}}^{(\mathcal{I})};Y).$$

Hence,

$$\begin{aligned} \lim_{N\to\infty}R_{dN} &= \lim_{N\to\infty}\frac{1}{N}I(S_{\mathcal{D}}^N;Y^N) \\ &\geq \sum_{\mathcal{I}\subset\mathcal{D}}(1-\beta)^{|\mathcal{I}|}\beta^{d-|\mathcal{I}|}I(X_{\mathcal{I}}^{(\mathcal{I})};Y). \end{aligned}$$

Optimizing over $p(s_{\mathcal{D}})$ concludes the proof of the achievability of Theorem 2.

ACKNOWLEDGMENT

This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and the National Science Foundation under Contract CCR-0311055.

REFERENCES

[1] L. Tong, Q. Zhao, and S. Adireddy, "Sensor Networks with Mobile Agents," in *Proc. 2003 Military Communications Intl Symp.*, Boston, MA, Oct. 2003.

[2] Z. Yang and L. Tong, "On the maximum achievable rate of information retrieval from cooperative sensor networks with misinformed sensors," submitted to *IEEE Intl. Symp. on Info. Theory, 2004.*