Opportunistic In-Network Computation for Wireless Sensor Networks

Sang-Woon Jeon
Department of ICE
Andong National University
Andong, South Korea
E-mail: swjeon@anu.ac.kr

Bang Chul Jung
Department of ICE
Gyeongsang National University
Tongyeong, South Korea
E-mail: bcjung@gnu.ac.kr

***Abstract*—Function computation over wireless sensor networks is investigated, where K sensors measure their observations and a fusion center wishes to estimate a pre-defined function of the observations via *fading* multiple access channels (MACs). The arithmetic sum and type functions are considered since they yield various fundamental sample statistics such as mean, variance, maximum, minimum, etc. We propose a novel *opportunistic* in-network computation (INC) scheme in which a subset of sensors with large channel gains opportunistically participate in the transmission at each time slot, while all sensors in a network simultaneously send their observations or only a single sensor sends its observation in the conventional INC schemes. We analyze the *ergodic* computation rate of the proposed INC scheme and prove that it achieves a non-vanishing computation rate even when the number of sensors K tends to infinity, which provides a significant rate improvement compared to the conventional INC schemes whose computation rates converge to zero as K increases.**

***Index Terms*—In-network computation, fading channels, lattice codes, opportunistic communication, wireless sensor networks.**

I. Introduction

Contrary to traditional wireless networks, the main goal of communications in wireless sensor networks (WSNs) is to compute some pre-defined *functions* of sensor observations (also called sensor readings) at a fusion center, rather than obtaining the observations themselves [1]. Applications of WSNs include disaster alarm, environmental monitoring, etc. For example, many sensor applications involve the sample mean, e.g., the average temperature from several temperature readings.

Unlike point-to-point channels, designing source and channel coding separately is quite suboptimal for function computing over general WSNs, especially when the network size increases. To overcome such limitation of the source–channel separation approach, communication techniques considering a joint design of source and channel coding have been actively studied in the literature [2]–[10], which is referred to as *joint source–channel coding*. The potential of linear source coding has been captured in [4], applying the linear source coding in [2] for the function computation over Gaussian multiple access channels (MACs). An efficient way of computing the modulo sum or the sum of Gaussian sources over Gaussian MACs using lattice codes has been proposed in [4], [6]. This lattice-based computation has been recently extended to multiple receivers called compute-and-forward [5], in which each relay computes or decodes linear combination of the sources. More recently, linear source coding and lattice-based computation have been applied for computing the arithmetic sum and type functions in [9]. Interactive communication between sensors in order to efficiently compute the type-threshold function has been studied in [7], [10]. In spite of the previous studies, however, relatively little progress has been made so far on how to efficiently compute functions under fading environment.

In this paper, therefore, we study the *function computation problem over the fading MAC*, which has been served as a fundamental building block for general WSNs. We propose a *opportunistic* INC framework by considering the time-varying nature of fading channels, in which a subset of sensors with large channel gains opportunistically participate in the transmission at each time slot. We further prove that the opportunistic INC framework achieves a non-vanishing computation rate even when the number of sensors in the network tends to infinity, which significantly improves the previous computation rates converging to zero as the number of sensors increases.

Notations: Let $[1:n] := \{1, 2, \cdots, n\}$, $\mathsf{C}(x) := \log(1+x)$, and $\mathsf{C}^{+}(x) := \max\{\log(x), 0\}$. Let $\operatorname{diag}(\{a_i\}_{i\in[1:n]})$ denote the diagonal matrix consisting of a_1 to a_n as its diagonal elements, $\mathbf{1}_{(\cdot)}$ denote the indicator function of an event, and $\operatorname{card}(\cdot)$ denote the cardinality of a set. For a random variable A, $H(A)$ denotes the entropy of A.

II. Problem Formulation

Consider the computation over the fading MAC depicted in Fig. 1, in which the fusion center wishes to compute a desired function of K sources observed by each of K sensor nodes. In particular, sensor $i \in [1:K]$ is assumed to observe a length-k discrete source vector $[s_i[1], \cdots, s_i[k]] \in [1:p]^k$ and the fusion center computes $f(s_1[j], s_2[j], \cdots, s_K[j])$ for all $j \in [1:k]$, where $p \in \mathbb{N}$ denotes the number of sample values. For convenience, denote the jth set of K sources as $\mathbf{s}[j] = [s_1[j], s_2[j], \cdots, s_K[j]]$, where $j \in [1:k]$. At each $j \in [1:k]$, $\mathbf{s}[j]$ is assumed to be independently drawn from a joint probability mass function $p_{\mathbf{S}}(\cdot)$.

The received signal at the fusion center at time slot t is given by

$$y[t] = \sum_{i=1}^{K} h_i[t]x_i[t] + z[t], \qquad (1)$$

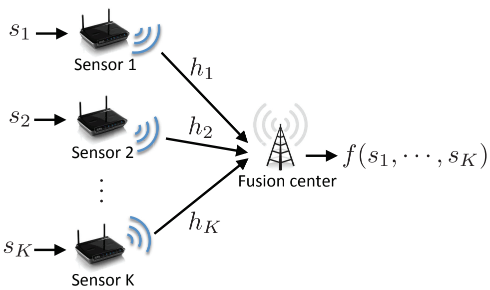

Fig. 1. INC model over fading MAC

where $x_i[t]$ denotes the transmit signal of sensor i at time slot t and $h_i[t]$ denotes the wireless channel coefficient from sensor i to the fusion center at time slot t. $z[t]$ represents the additive Gaussian noise at time slot t, assumed to follow $\mathcal{CN}(0,1)$ and be independent over time slots. We assume time-varying channel coefficients in which $h_i[t]$ is i.i.d. drawn from a continuous distribution $f_h(\cdot)$ for each time slot. We further assume that global channel state information is available at all sensors and the fusion center. All sensors are assumed to have an identical average power constraint, i.e., $\mathsf{E}[|x_i[t]|^2] \le P$ for all $i \in [1:K]$. In this paper, we focus on two types of desired functions: the arithmetic sum function and the type (or frequency histogram) function, as considered in [3], [9].

For completeness, we state the formal definition of the desired function in the following.

Definition 1 (Desired function): Let $\mathbf{s} = [s_1, \cdots, s_K] \in [1:p]^K$. Then $f(\mathbf{s}) = \{\sum_{i=1}^K a_{li} s_i\}_{l=1}^L$ and $f(\mathbf{s}) = \{b_1(\mathbf{s}), \cdots, b_p(\mathbf{s})\}$ for the arithmetic sum computation and the type computation respectively, where $a_{li} \in [0:q]$ and $b_m(\mathbf{s}) = \sum_{i=1}^K \mathbf{1}_{s_i=m}$. $\Diamond$

Note that the desired function in Definition 1 is locally computable, which will be exploited in the proposed INC scheme. The definition of a locally computable function is given in the following.

Definition 2 (Locally computable function): Suppose that $\{\lambda_i\}_{i=1}^N$ are N partition sets of $[1:K]$, i.e., $\lambda_i \cap \lambda_j = \emptyset$ for all $i,j \in [1:N]$ with $i \ne j$ and $\bigcup_{i=1}^N \lambda_i = [1:K]$. A function is said to be locally computable if there exists $g(\cdot)$ for any $\{\lambda_i\}_{i=1}^N$ satisfying $f(\mathbf{s}) = g(f(\{s_i\}_{i\in\lambda_1}), \cdots, f(\{s_i\}_{i\in\lambda_N}))$ $\Diamond$.

Let $\mathbf{S} = [S_1, \cdots, S_K] \in [1:p]^K$ be the random source vector associated with $p_{\mathbf{S}}(\cdot)$. Then $f(\mathbf{S})$ is the desired function induced by the random source vector $\mathbf{S}$. Denote $U_l = \sum_{i=1}^K a_{li} S_i$ for $l \in [1:L]$ and $B_m = \sum_{i=1}^K \mathbf{1}_{S_i=m}$ for $m \in [1:p]$. Then $f(\mathbf{S}) = (U_1, \cdots, U_L)$ for the arithmetic sum function and $f(\mathbf{S}) = (B_1, \cdots, B_p)$ for the type function.

Definition 3 (Computation rate): The computation rate $R := \lim_{n\to\infty} \frac{k(n)}{n} H(f(\mathbf{S}))$ is said to be achievable if there exists a sequence of length-n block codes such that $\Pr\left[\bigcup_{j=1}^k \hat{f}(\mathbf{s}[j]) \ne f(\mathbf{s}[j])\right] \to 0$ as n increases. $\Diamond$

From Definition 3, the computation rate R bits/sec/Hz is the number of information bits with respect to the desired function delivered by each channel use.

III. Preliminaries and Main Results

In this section, we briefly introduce existing INC schemes and their limitation to fading environment, and then show our main results. For easy presentation, we omit the time index t in the rate expressions.

A. Previous Work

In [9], computing the arithmetic sum and type functions, defined in Definition 1, has been studied for Gaussian (non-fading) MAC, i.e., $h_i[t] = h_i$ for all t. The authors showed that $R(h_1, \cdots, h_K) = \mathsf{C}^+\left(\frac{1}{K} + \min_{i\in[1:K]} |h_i|^2 P\right)$ is achievable, see [9, Theorem 3]. Therefore, it can be shown that $R = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \min_{i\in[1:K]} |h_i|^2 P\right)\right]$ is achievable by applying the result of [9, Theorem 3] to the fading MAC in Section II. To improve this computation rate for fading environment, long-term power control has been considered in [9, Theorem 5], provided that

$$R = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \frac{\min_{i\in[1:K]} |h_i|^2 P}{\mathsf{E}[\min_{i\in[1:K]} |h_i|^2/|h_1|^2]}\right)\right] \quad (2)$$

is achievable for the fading MAC in Section II, where $\frac{1}{\mathsf{E}[\min_{i\in[1:K]} |h_i|^2/|h_1|^2]} \ge 1$ represents the gain from the long-term power control. However, for i.i.d. Rayleigh fading channels, i.e., $h_i[t] \sim \mathcal{CN}(0,1)$, the computation rate in (2) converges to zero as K increases. Another approach is for each sensor to transmit separately in each time slot (without INC), achieving the computation rate of $R = \frac{1}{K}\mathsf{E}\left[\mathsf{C}(|h_1|^2 P)\right]$, which also converges to zero as K increases due to the term $1/K$. To the best of our knowledge, the computation rates achievable by directly applying the conventional INC schemes for the considered fading MAC decrease as the number of sensors K increases and eventually converge to zero in the limit of large K [3]–[5], [8], [9].

B. Main Results

We derive the computation rate of the proposed opportunistic INC framework over the fading MAC in Theorem 1 and prove that it achieves a non-vanishing computation rate regardless of K in Corollary 1.

Theorem 1: For any $M, N \in \mathbb{N}$ satisfying $MN = K$, the computation rate of the proposed opportunistic INC over the fading MAC described in Section II is given by

$$R = \frac{1}{N}\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 KP}{\sum_{i=1}^M \mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_i}|^2}\right]}\right)\right], \quad (3)$$

where $\{\pi_i\}_{i=1}^K$ denotes the set of ordered sensor indices of $[1:K]$ such that $|h_{\pi_1}| \ge |h_{\pi_2}| \ge \cdots \ge |h_{\pi_K}|$.

Proof: We refer to Section IV for the proof. ■

The proposed INC framework exploits both *the superposition property of wireless channels*, which has been used for in-network computing in non-fading networks [4], [5], [8], [9], and *the locally computable property of the desired function*, which has been used for computing in tree networks [3] and interactive computing between nodes [7], [10]. For instance, Theorem 1 attains the result in [9, Theorem 5] by setting $M = K$ and $N = 1$ and also attains the communication-based computation by seting $M = 1$ and $N = K$.

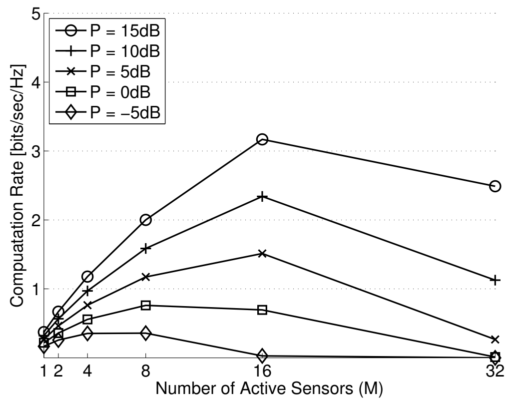

Fig. 2. Computation rates with respect to M when $K = 32$ for i.i.d. Rayleigh fading channels.

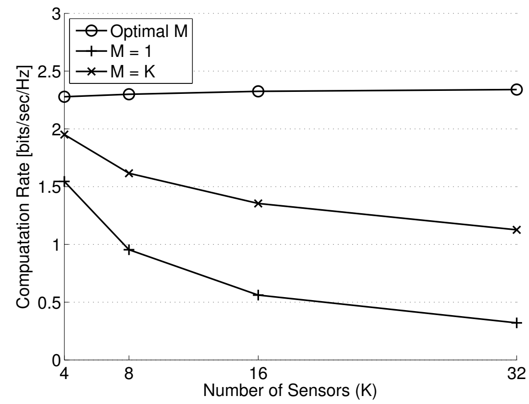

Fig. 3. Computation rates with respect to K when $P = 10$ dB for i.i.d. Rayleigh fading channels.

Corollary 1: As K increases, the computation rate achievable by Theorem 1 is given by

$$R = \min\left(\Delta\,\mathsf{E}\left[\mathsf{C}\left(|h_1|^2 P\right)\right], \frac{1-\Delta}{2}\mathsf{E}\left[\mathsf{C}^+(2\mu P)\right]\right), \quad (4)$$

where $\Delta \in (0,1)$ and μ denotes the median of the distribution $f_{|h|^2}(\cdot)$, which is induced by $f_h(\cdot)$.

Proof: We refer to the full paper in preparation for the proof. ■

Notice that Δ and μ in Corollary 1 are not a function of K and, as a consequence, R in Corollary 1 is not a function of K. Therefore, the proposed scheme achieves a non-vanishing computation rate if $P > \frac{1}{2\mu}$. For i.i.d. Rayleigh fading, for instance, $f_{|h|^2}(\cdot)$ follows the exponential distribution with parameter one, i.e., $f_{|h|^2}(x) = \exp(-x)$ and $\mu = \ln(2)$. Hence a non-vanishing computation rate is achievable if $P > \frac{1}{2\ln(2)}$, which is approximately given by -1.4 dB.

Fig. 2 plots the computation rate of the proposed opportunistic INC in Theorem 1 with respect to M when $K = 32$. The results demonstrate that the computation rate in Theorem 1 with optimally chosen M outperforms the existing INC schemes, which are the cases where $M = 1$ and $M = K$ in the figure. Fig. 3 shows the computation rate of the opportunistic INC with optimal M as the number of sensors K increases. As proved in Corollary 1, the opportunistic INC with optimal M achieves a non-vanishing computation rate even as K increases, while the computation rates attained by the conventional INC schemes converge to zero as K increases.

IV. Opportunistic In-Network Computation

In this section, we prove Theorem 1. For each time slot $t \in [1:n]$, let us define $\{\pi_i[t] \in [1:K]\}_{i=1}^K$ as the set of reordered sensor indices such that $|h_{\pi_1[t]}[t]| \geq |h_{\pi_2[t]}[t]| \geq \cdots \geq |h_{\pi_K[t]}[t]|$.

A. Opportunistic Local Function Computation

At each time slot t, the M sensors in $\{\pi_i[t]\}_{i=1}^M$ participate in the transmission and the fusion center computes the local function $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$. Let $R'(t)$ denote the computation rate at time slot t for the local function $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$ and $P_i[t]$ denote the transmit signal power of sensor i at time slot t, where $P_{\pi_i[t]}[t] = 0$ for $i \in [M+1:K]$. We apply the same computing code proposed in [9] for each local function computation, see [9, Theorem 3] for the detailed code construction. From [9, Theorem 3],

$$R'[t] = \mathsf{C}^+\left(\frac{1}{M} + \min_{i\in[1:M]} |h_{\pi_i[t]}[t]|^2 P_{\pi_i[t]}[t]\right) \quad (5)$$

is achievable for the local function computation of $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$. Then by setting

$$P_{\pi_i[t]}[t] = \frac{KP}{\sum_{j=1}^M \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]}\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_i[t]}[t]|^2} \text{ for } i \in [1:M], \quad (6)$$

which satisfies the average power constraint P, and applying the same coding strategy over large enough time slots satisfying $\{\pi_i[t]\}_{i=1}^M = \{\pi_i\}_{i=1}^M$, the ergodic computation rate of

$$R' := \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 KP}{\sum_{i=1}^M \mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_i}|^2}\right]}\right)\right] \quad (7)$$

is achievable for computing $f(s_{\pi_1}, \cdots, s_{\pi_M})$. In the following, we state how to apply such local computing to attain sample-by-sample desired functions $\{f(\mathbf{s}[j])\}_{j\in[1:k]}$.

B. Construction of the Desired Function

Since only M sensors opportunistically participate in the transmission at each time slot, in order to construct the desired function $f(\mathbf{s})$ in Definition 1, the fusion center needs N local functions. Let λ_1 to λ_N denote sensor subsets each consisting of M sensors. Then, from the locally computable property in Definition 2, the fusion center is able to construct $f(\mathbf{s})$ by using the computed $f(\{s_i\}_{i\in\lambda_1})$ to $f(\{s_i\}_{i\in\lambda_N})$ if $\bigcup_{l=1}^N \lambda_l = [1:K]$ is satisfied. In order to exploit this property, however, the sample indices of the N local functions should be the same. We explain how to obtain such sample-by-sample desired functions in the following.

Let $\Lambda = \{\lambda \subseteq [1:K] : \mathrm{card}(\lambda) = M\}$ denote the set of all sensor subsets consisting of M sensors in each subset, where $\mathrm{card}(\Lambda) = \binom{K}{M}$. For $\lambda \in \Lambda$, define $\mathcal{T}_\lambda = \{t \in [1:n]:$

TABLE I. The transmitted source at each sensor and the computed local function at the fusion center.

	sensor 1	sensor 2	sensor 3	sensor 4	fusion center
$t=1$	$s_1[2]$	$\emptyset$	$s_3[2]$	$\emptyset$	$f(s_1[2],s_3[2])$
$t=2$	$\emptyset$	$s_2[3]$	$s_3[3]$	$\emptyset$	$f(s_2[3],s_3[3])$
$t=3$	$\emptyset$	$\emptyset$	$s_3[1]$	$s_4[1]$	$f(s_3[1],s_4[1])$
$t=4$	$\emptyset$	$s_2[2]$	$\emptyset$	$s_4[2]$	$f(s_2[2],s_4[2])$
$t=5$	$s_1[1]$	$s_2[1]$	$\emptyset$	$\emptyset$	$f(s_1[1],s_2[1])$
$t=6$	$\emptyset$	$s_2[4]$	$s_3[4]$	$\emptyset$	$f(s_2[4],s_3[4])$
$t=7$	$\emptyset$	$s_2[5]$	$\emptyset$	$s_4[5]$	$f(s_2[5],s_4[5])$
$t=8$	$s_1[6]$	$s_2[6]$	$\emptyset$	$\emptyset$	$f(s_1[6],s_2[6])$
$t=9$	$s_1[3]$	$\emptyset$	$\emptyset$	$s_4[3]$	$f(s_1[3],s_4[3])$
$t=10$	$s_1[5]$	$\emptyset$	$s_3[5]$	$\emptyset$	$f(s_1[5],s_3[5])$
$t=11$	$\emptyset$	$\emptyset$	$s_3[6]$	$s_4[6]$	$f(s_3[6],s_4[6])$
$t=12$	$s_1[4]$	$\emptyset$	$\emptyset$	$s_4[4]$	$f(s_1[4],s_4[4])$

$\{\pi_i[t]\}_{i=1}^M=\lambda\}$ as the set of time slot indices that the sensors in λ participate in the transmission. We further define

$$\Omega=\left\{\omega=(\lambda_1,\cdots,\lambda_N)\in\Lambda^N:\bigcup_{l=1}^{N}\lambda_l=[1:K]\right\} \tag{8}$$

as the set of all possible N sensor subsets that can be used for constructing the desired function from the locally computable property in Definition 2, where $\mathrm{card}(\Omega)=\prod_{l=0}^{N-1}\binom{K-Ml}{M}$. For $\lambda\in\Lambda$, let $\Omega_\lambda=\{\omega\in\Omega:\lambda\in\omega\}$ as the set of N sensor subsets that include λ as an element, where $\mathrm{card}(\Omega_\lambda)=N\prod_{l=1}^{N-1}\binom{K-Ml}{M}$.

1) Sample-by-sample computing: We first provide an intuitive explanation on how to obtain sample-by-sample desired functions based on the case where $K=4$ and $M=N=2$. For this case,

$$\begin{aligned}\Lambda&=\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\},\\ \Omega&=\{((1,2),(3,4)),((1,3),(2,4)),((1,4),(2,3)),\\ &\qquad((2,3),(1,4)),((2,4),(1,3)),((3,4),(1,2))\}.\end{aligned} \tag{9}$$

From Ω, we are able to define the transmission scheme at each sensor and the desired function computation by using computed local functions at the fusion center. Suppose that $n=12$ and $\mathcal{T}_{(1,2)}=\{5,8\}$, $\mathcal{T}_{(1,3)}=\{1,10\}$, $\mathcal{T}_{(1,4)}=\{9,12\}$, $\mathcal{T}_{(2,3)}=\{2,6\}$, $\mathcal{T}_{(2,4)}=\{4,7\}$, $\mathcal{T}_{(3,4)}=\{3,11\}$. Then, the transmission of each sensor and the local function computation at the fusion center are given in Table I. For simplicity, we assume that the number of local functions computable by a single channel use is equal to one. Specifically, at $t=1$, sensors 1 and 3 send $s_1[2]$ and $s_3[2]$ respectively and the fusion center computes $f(s_1[2],s_3[2])$. Note that they send their second sources at $t=1$ since $(1,3)$ firstly appears in the second element in Ω. Similarly, at $t=2$, sensors 2 and 3 send $s_2[3]$ and $s_3[3]$ respectively and the fusion center computes $f(s_2[3],s_3[3])$ since $(2,3)$ firstly appears in the third element in Ω. Each sensor and the fusion center perform the same procedure for the rest of time slots. For instance, at $t=10$, sensors 1 and 3 send $s_1[5]$ and $s_3[5]$ respectively since $(1,3)$ secondly appears in the fifth element in Ω.

After computing 12 local functions as in Table I, the fusion center attains 6 desired functions again based on Ω. Specifically, the first element in Ω is given by $((1,2),(3,4))$ and, therefore, the two local functions computed at time slots 5 and 3 are used to construct $f(\mathbf{s}[1])$, i.e., $f(\mathbf{s}[1])=g(f(s_1[1],s_2[1]),f(s_3[1],s_4[1]))$. Similarly, $f(s_1[2],s_3[2])$ and $f(s_2[2],s_4[2])$ computed at time slots 1 and 4 respectively are used to construct $f(\mathbf{s}[2])$ from the second element in Ω. In the same manner, the fusion center attains the rest of the desired functions.

2) Detailed construction: In the above example, we assume that $\mathrm{card}(\mathcal{T}_\lambda)$ is the same for all $\lambda\in\Lambda$. In practice, however, $\mathrm{card}(\mathcal{T}_\lambda)$ is random, varying over channel realizations. The following lemma characterizes the minimum deterministic number of $\mathrm{card}(\mathcal{T}_\lambda)$, which is the same for all $\lambda\in\Lambda$, in the limit of large n.

Lemma 1: For any $\epsilon>0$, the probability that

$$\left|\frac{1}{n}\mathrm{card}(\mathcal{T}_\lambda)-\frac{1}{\binom{K}{M}}\right|\geq\epsilon \tag{10}$$

for all $\lambda\in\Lambda$ is lower bounded by $1-\frac{\binom{K}{M}}{4n\epsilon^2}$.

Proof: Since channel coefficients are i.i.d., $\Pr(\{\pi_i[t]\}_{i=1}^M=\lambda)$ is the same for all $\lambda\in\Lambda$, given by $\Pr(\{\pi_i[t]\}_{i=1}^M=\lambda)=1/\binom{K}{M}$ for all $\lambda\in\Lambda$. Therefore, from the strong typicality in [11, Lemma 2.12], Lemma 1 holds. ■

By setting $\epsilon=\frac{1}{\log n}$ in Lemma 1, $\mathrm{card}(\mathcal{T}_\lambda)\geq\frac{n}{\binom{K}{M}}-\frac{n}{\log n}$ for all $\lambda\in\Lambda$ with probability greater than $1-\frac{\binom{K}{M}(\log n)^2}{4n}$, which converges to one as n increases. Therefore, from Lemma 1, at least

$$T:=\frac{\frac{n}{\binom{K}{M}}-\frac{n}{\log n}}{\mathrm{card}(\Omega_\lambda)}=\frac{\frac{n}{\binom{K}{M}}-\frac{n}{\log n}}{N\prod_{l=1}^{N-1}\binom{K-Ml}{M}} \tag{11}$$

time slots in $\mathcal{T}_\lambda$ can be used for computing the desired function based on Ω. Denote such T time slots in $\mathcal{T}_\lambda$ as $\mathcal{T}_{\lambda,\omega}$, where $\mathcal{T}_{\lambda,\omega_1}\cap\mathcal{T}_{\lambda,\omega_2}=\emptyset$ for $\omega_1\neq\omega_2\in\Omega_\lambda$. Specifically, for given $\omega\in\Omega$, the time slots in $\mathcal{T}_{\lambda,\omega}$ are used to compute the local function $f(\{s_i\}_{i\in\lambda})$ for all $\omega\in\Omega_\lambda$, see also the example in Section IV-B1 and Table I.

Let $\mathbf{x}_i(\lambda,\omega)\in\mathbb{C}^{T\times1}$ denote the length-T time-extended transmit signal vector of sensor i during $t\in\mathcal{T}_{\lambda,\omega}$. Specifically, we construct

$$\mathbf{x}_i(\lambda,\omega)=\mathbf{\Gamma}_i(\lambda,\omega)\mathbf{x}'_i(\lambda,\omega), \tag{12}$$

where $\mathbf{\Gamma}_i(\lambda,\omega)=\mathrm{diag}\left(\left\{\sqrt{P_i[t]}\frac{|h_i[t]|}{h_i[t]}\right\}_{t\in\mathcal{T}_{\lambda,\omega}}\right)$ and $\mathbf{x}'_i(\lambda,\omega)$ is the lattice-based transmit signal vector for the distributed INC satisfying the average transmit power of one used in [9, Theorem 3]. Here, from Section IV-A,

$$P_i[t] = \begin{cases} \frac{KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]} \frac{|h_{\pi_M[t]}[t]|^2}{|h_i[t]|^2} & \text{for } i \in \{\pi_j[t]\}_{j=1}^{M}, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Obviously, $\mathbf{x}_i(\lambda, \omega) = \mathbf{0}$ from (13) if $i \notin \lambda$ since the M sensors with the largest channel gains are in λ for $t \in \mathcal{T}_{\lambda,\omega}$. Then, for all $\lambda \in \Lambda$ and $\omega \in \Omega_\lambda$, sensor i transmits $\mathbf{x}_i(\lambda, \omega)$ during $t \in \mathcal{T}_{\lambda,\omega}$.

Let $\mathbf{y}(\lambda, \omega) \in \mathbb{C}^{T \times 1}$ denote the length-T time-extended received signal vector of the fusion center during $t \in \mathcal{T}_{\lambda,\omega}$, that is given by

$$\mathbf{y}(\lambda, \omega) = \sum_{i \in \lambda} \mathbf{H}_i(\lambda, \omega)\mathbf{x}_i(\lambda, \omega) + \mathbf{z}(\lambda, \omega), \tag{14}$$

where $\mathbf{H}_i(\lambda, \omega) = \mathrm{diag}(\{h_i[t]\}_{t \in \mathcal{T}_{\lambda,\omega}})$ and $\mathbf{z}(\lambda, \omega)$ is the length-T time-extended noise vector during $t \in \mathcal{T}_{\lambda,\omega}$. Then, from (12),

$$\begin{aligned} \mathbf{y}(\lambda, \omega) &= \sum_{i \in \lambda} \mathbf{H}_i(\lambda, \omega)\mathbf{\Gamma}_i(\lambda, \omega)\mathbf{x}'_i(\lambda, \omega) + \mathbf{z}(\lambda, \omega) \\ &= \sum_{i \in \lambda} \mathbf{H}'_i(\lambda, \omega)\mathbf{x}'_i(\lambda, \omega) + \mathbf{z}(\lambda, \omega), \end{aligned} \tag{15}$$

where $\mathbf{H}'_i(\lambda, \omega) = \mathrm{diag}(\{h'_i[t]\}_{t \in \mathcal{T}_{\lambda,\omega}})$ and

$$\begin{aligned} h'_i[t] &= h_i[t]\sqrt{P_i[t]}\frac{|h_i[t]|}{h_i[t]} \\ &= \sqrt{\frac{KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]}}|h_{\pi_M[t]}[t]|. \end{aligned} \tag{16}$$

Therefore, from (7), the computation rate of

$$R' = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_j}|^2}\right]}\right)\right] \tag{17}$$

is achievable for computing the local function $f(\{s_i\}_{i \in \lambda})$. More specifically, the fusion center computes $\{f(\{s_i[l]\}_{i \in \lambda})\}_{l \in [1:R'T/H(f(\mathbf{S}))]}$ during $t \in \mathcal{T}_{\lambda,\omega}$. We refer to the example in Section IV-B1 for the source rearrangement at each sensor in order to compute sample-by-sample local functions.

Hence, during $t \in \bigcup_{\lambda \in \Lambda : \lambda \in \omega} \mathcal{T}_{\lambda,\omega}$, the fusion center computes $\{f(\{s_i[l]\}_{i \in \lambda})\}_{l \in [1:R'T/H(f(\mathbf{S}))]}$ for all $\lambda \in \omega$. By using the locally computable property of the desired function in Definition 2, the fusion center then computes the desired functions $\{f(s_1[l], \cdots, s_K[l])\}_{l \in [1:R'T/H(f(\mathbf{S}))]}$. Since $\mathrm{card}(\Omega) = \prod_{l=0}^{N-1} \binom{K-Ml}{M}$, the number of the computed desired functions during $t \in [1:n]$ is given by

$$k = \left(\prod_{l=0}^{N-1} \binom{K-Ml}{M}\right) \frac{R'T}{H(f(\mathbf{S}))} \tag{18}$$

and, as a result, the computation rate of

$$\begin{aligned} R &= \frac{kH(f(\mathbf{S}))}{n} \\ &\overset{(a)}{=} \frac{\left(\prod_{l=0}^{N-1} \binom{K-Ml}{M}\right) R'T}{n} \\ &\overset{(b)}{=} \frac{\left(\prod_{l=0}^{N-1} \binom{K-Ml}{M}\right) R'}{n} \frac{\frac{n}{\binom{K}{M}} - \frac{n}{\log n}}{N \prod_{l=1}^{N-1} \binom{K-Ml}{M}} \\ &= \frac{1}{N} R' \left(1 - \frac{\binom{K}{M}}{\log n}\right) \end{aligned} \tag{19}$$

is achievable, where (a) and (b) follow from (18) and (11) respectively. In conclusion, (3) is achievable as n increases, which complete the proof of Theorem 1.

V. Conclusion

In this paper, we investigated the function computation problem in WSNs, focusing on an efficient INC strategy for fading environment. The proposed opportunistic INC framework exploits both the superposition property of wireless channel and the locally computable property of the desired function, combining with opportunistic transmission. We showed that a non-vanishing computation rate can be achieved by the opportunistic INC as the number of sensors in the network increases.

Acknowledgement

This work was supported by ICT R&D program of MSIP/IITP. [B0101-15-1272, Development of Device Collaborative Giga-Level Smart Cloudlet Technology]

References

[1] A. Giridhar and P. R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, pp. 98–107, Apr. 2006.

[2] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 219–221, Mar. 1979.

[3] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 755–764, Apr. 2005.

[4] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3498–3516, Oct. 2007.

[5] ——, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.

[6] R. Soundararajan and S. Vishwanath, "Communicating linear functions of correlated Gaussian sources over a MAC," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1853–1860, Mar. 2012.

[7] N. Ma, P. Ishwar, and P. Gupta, "Interactive source coding for function computation in collocated networks," *IEEE Trans. Inf. Theory*, vol. 58, pp. 4289–4305, Jul. 2012.

[8] J. Zhan, S. Y. Park, M. Gastpar, and A. Sahai, "Linear function computation in networks: Duality and constant gap results," *IEEE J. Select. Areas Commun.*, vol. 31, pp. 620–638, Apr. 2013.

[9] S.-W. Jeon, C.-Y. Wang, and M. Gastpar, "Computation over Gaussian networks with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 60, pp. 7841–7861, Dec. 2014.

[10] C.-Y. Wang, S.-W. Jeon, and M. Gastpar, "Interactive computation of type-threshold functions in collocated broadcast–superposition networks," in *arXiv:cs.IT/1310.2860*, Oct. 2013.

[11] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.