Opportunistic Function Computation for Wireless Sensor Networks

Sang-Woon Jeon, *Member, IEEE*, and Bang Chul Jung, *Senior Member, IEEE*

***Abstract*—Function computation over wireless sensor networks is investigated, where a set of K sensors observe their sensor readings and a fusion center wishes to learn a predefined function of the sensor readings via *fading* multiple access channels (MACs). In this paper, the arithmetic sum and type functions are considered since they can yield various fundamental sample statistics such as mean, variance, maximum, and minimum. We propose a novel *opportunistic* in-network computation (INC) framework in which a subset of sensors with large channel gains opportunistically participate in the transmission at each time, while all sensors simultaneously send their observations or only a single sensor sends its observation in the conventional schemes. We mathematically analyze the long-term average computation rate of the proposed INC, and prove that it achieves a nonvanishing computation rate even when the number of sensors K tends to infinity, which is in fact a significant improvement and the first theoretical result in fading MACs. Note that the computation rates of the conventional schemes become zero as K increases. We further show that a similar multiuser diversity gain is still achievable under delay constraints, which implies that the proposed INC is restricted to exploit a fixed and finite number of time slots (or fading instances) for the function computation.**

***Index Terms*—In-network computation, fading channels, lattice codes, opportunistic communication, power adaptation, 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. From the current digital communication paradigm, each sensor may compress its sensor readings, independently of channel coding, and then the compressed sensor information is reliably sent to the fusion center via an appropriate channel coding, which is referred to as *source–channel separation*. This source–channel separation approach is shown to be optimal for point-to-point channels [2]. However, designing source and channel coding separately is quite suboptimal for function computation over general WSNs, especially when the network size increases.

To overcome such limitations of the source–channel separation approach, communication techniques considering a joint design of source and channel coding have been actively studied in the literature [3]–[12], which is referred to as *joint source–channel coding*. The potential of linear source coding has been captured in [7], applying the linear source coding in [3] 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 [7], [9], see also [13]–[15] for lattice code construction. This lattice-based computation has been recently extended to multiple receivers called compute-and-forward [8], in which each relay computes or decodes linear combination of the sources. In [10], a similar lattice code construction has been used for computing a linear function over linear finite field networks and the sum of Gaussian sources over Gaussian networks. More recently, linear source coding and lattice-based computation have been applied for computing the arithmetic sum and type functions in [11]. Interactive communication between sensors in order to efficiently compute the type-threshold function has been studied in [16] from the source coding perspective and in [12] from the joint source–channel coding perspective.

Along with the above information-theoretical studies, a simple computation strategy, called analog function computation (AFC), has been considered for wireless MACs [17]–[19]. The AFC adopts both pre-processing at each sensor and post-processing at a fusion center for computing or estimating functions, in which all sensors concurrently participate in the transmission and the channel inversion technique is used at each sensor in order to deal with fading phenomenon [17]. The AFC was further analyzed in terms of the maximum achievable computation rate in clustered sensor networks [18], but fading and noise effects were not considered.

The underlying principle of the most joint source–channel coding designs for function computing is to exploit a similarity between the superposition properties of wireless channels and desired functions. We can easily find the examples that the joint source–channel coding approach improves an achievable computation rate, compared with the source–channel separation

Manuscript received February 11, 2015; revised July 23, 2015 and November 16, 2015; accepted February 9, 2016. Date of publication February 23, 2016; date of current version June 7, 2016. This work was supported by the ICT R&D program of MSIP/IITP (1391104004, Development of Device Collaborative Giga-Level Smart Cloudlet Technology). This paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, June 2015. The associate editor coordinating the review of this paper and approving it for publication was M. Vu. *(Corresponding author: Bang Chul Jung.)*

S.-W. Jeon is with the Department of Information and Communication Engineering, Andong National University, Andong, South Korea (e-mail: swjeon@anu.ac.kr).

B. C. Jung is with the Department of Electronics Engineering, Chungnam National University, Daejeon, South Korea (e-mail: bangchuljung@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2016.2533379

approach. In spite of the previous studies, however, relatively little progress has been made so far on how to efficiently compute functions under fading environment. For instance, consider the computation of the arithmetic sum (or sample mean) of K sources each of which is observed by K sensors. Suppose that the K sensors are connected with a fusion center via the fading MAC, i.e., the channel coefficients are i.i.d. and vary independently over time. To the best of our knowledge, the computation rates achievable by straightforwardly applying the current computation techniques decreases as the number of sensors K increases and eventually converges to zero in the limit of large K due to fading.

In this paper, we study the *function computation problem over the fading MAC*, which has served as a fundamental building block for general WSNs. We mainly focus on the *type function* computation that is able to yield various important sample statistics such as mean, variance, maximum, minimum, mode, median and so on, as mentioned in [20]. We propose a novel *opportunistic* in-network computation (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, while all sensors simultaneously send their observations or only a single sensor sends its observation in the conventional schemes. The proposed INC prohibits the computation rate from being limited by the sensors with deep fading by exploiting the opportunism of fading channels. We analyze the long-term *ergodic* computation rate of the proposed INC, achievable by applying it over large enough fading instances. We further prove that the proposed INC attains 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. We also show that a similar multi-user diversity (or opportunistic scheduling) gain from multiple sensors is still achievable when the proposed INC is restricted to exploit a fixed and finite number of fading instances.

The paper is organized as follows. Section II introduces the system model and formulates the problem. In Section III, we summarize the main results of this paper and compare them with the existing results. Section IV presents the proposed opportunistic INC framework in detail and analyzes the computation rate of the proposed INC. Section V presents the opportunistic INC framework under delay constraints. Numerical examples are provided in Section VI and extensions to more general networks are given in Section VII. Finally, concluding remarks are given in Section VIII.

II. System Model and Problem Statement

In this section, we introduce the wireless in-network computing model for WSNs and define its achievable computation rate. Throughout the paper, we denote $[1:n] := \{1, 2, \cdots, n\}$, $\mathsf{C}(x) := \log(1+x)$, and $\mathsf{C}^+(x) := \max\{\log(x), 0\}$. Let $\mathrm{diag}(\{a_i\}_{i\in[1:n]})$ 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 $\mathrm{card}(\cdot)$ denote the cardinality of a set. For a random variable A, $H(A)$ denotes the entropy of A.

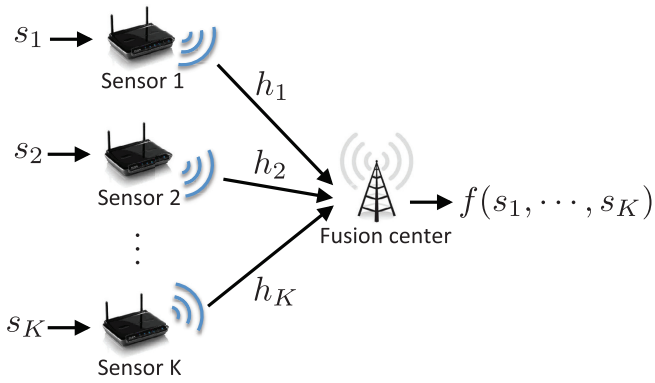


Fig. 1. Wireless in-network computing model.

A. Fading Multiple Access Channels

Consider the computation over the fading MAC depicted in Fig. 1, in which the fusion center is supposed to compute a desired function of K sources observed by each of K sensor nodes. The ith sensor node ($i \in [1:K]$) is assumed to observe a *discrete* source vector whose length is equal to k, i.e., $\mathbf{s}_i = [s_i[1], \cdots, s_i[k]] \in [1:p]^k$, where $p \in \mathbb{N}$ denotes the number of possible sample values.[1] The fusion center computes a sample-by-sample function of K sources, i.e.,

$$f(s_1[j], s_2[j], \cdots, s_K[j]) \text{ for all } j \in [1:k]. \tag{1}$$

For convenience, let the jth sample set of K sources as $\mathbf{s}[j] = [s_1[j], s_2[j], \cdots, s_K[j]]$, where $j \in [1:k]$.

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

$$y[t] = \sum_{i=1}^{K} h_i[t]x_i[t] + z[t], \tag{2}$$

where $x_i[t]$ represents the transmit signal of the ith sensor node at the tth time slot, $h_i[t]$ represents a complex-valued wireless channel from the ith sensor node to the fusion center at the tth time slot, and $z[t]$ denotes the complex Gaussian noise at the tth time slot, assumed to follow $\mathcal{CN}(0,1)$ and independent over time slots. We assume block fading channels in which $h_i[t]$ is i.i.d. drawn from a continuous distribution $f_h(\cdot)$ for each time slot but it remains unchanged during a given time slot. We further assume that global channel state information (CSI) is available at all sensor nodes and fusion center, i.e., $\{h_i[t]\}_{i=1}^K$ is available at time t.[2] All sensor nodes are assumed to have an identical average power constraint, i.e., $\mathsf{E}[|x_i[t]|^2] \le P$ for all $i \in [1:K]$.

Remark 1: In the considered model, i.i.d. fading is assumed, i.e., channel statistics are identical for all sensors, which makes asymptotic analysis tractable by applying the same opportunistic transmission strategy at each sensor in function computing.

[1]The measurement of the ith sensor node can be regarded to be sampled k times and each of the samples can be regarded to be mapped to one of numbers from 1 to p via a kind of quantization process.

[2]As we will explain in Section IV-A, the proposed scheme requires more relaxed channel information allowing an efficient implementation of the required channel feedback.

We will briefly consider how to extend to asymmetric networks suffering from different signal attenuation from each sensor to the fusion center in Section VII-B. ◇

In this paper, we focus on two desired functions: the arithmetic sum function and the type (or frequency histogram) function, as considered in [11], [12], [20]. For completeness, we formally define the desired function in the following.

Definition 1 (Desired function): Let $\mathbf{s} = [s_1, \cdots, s_K] \in [1 : p]^K$. For the arithmetic sum computation, the desired function is given by $f(\mathbf{s}) = \{\sum_{i=1}^K a_{li} s_i\}_{l=1}^L$, where $a_{li} \in [0 : q]$ and $q, L \in \mathbb{N}$. For the type computation, the desired function is given by $f(\mathbf{s}) = \{b_1(\mathbf{s}), \cdots, b_p(\mathbf{s})\}$, where $b_l(\mathbf{s}) = \sum_{i=1}^K \mathbf{1}_{s_i=l}$ for $l \in [1 : p]$.

The arithmetic sum function in this paper is defined as multiple weighted arithmetic sums, which includes the sample mean function $f(\mathbf{s}) = \frac{1}{K}\sum_{i=1}^K s_i$ and the case of estimating each of the K samples $f(\mathbf{s}) = \{s_l\}_{l=1}^K$ as special cases. Computing the type function is very powerful since any *symmetric function*[3] can be attained from the type function [11], [12], [20]. Most fundamental sample statistics such as mean, variance, maximum, minimum, median, and mode are symmetric functions. As seen in Definition 1, the type function is represented as multiple weighted arithmetic sums of indicator functions.

The proposed opportunistic INC constructs the desired function from a set of local functions by using the locally computable property of the desired function, which is defined 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 $i \neq j$ 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 $f(\mathbf{s}) = g(f(\{s_i\}_{i\in\lambda_1}), \cdots, f(\{s_i\}_{i\in\lambda_N}))$ ◇

Clearly, both arithmetic sum function and type function are locally computable from Definition 2.

In this paper, we assume arbitrarily correlated stationary and ergodic sources. The following definition formally states the underlying probability distribution and the corresponding random variables regarding the set of K sources.

Definition 3 (Sources): Let $\mathbf{S} = [S_1, \cdots, S_K] \in [1 : p]^K$ be a random vector associated with a joint probability mass function $p_{\mathbf{S}}(\cdot)$. At each $j \in [1 : k]$, $\mathbf{s}[j]$ is assumed to be independently drawn from $p_{\mathbf{S}}(\cdot)$. ◇

Let $f(\mathbf{S})$ denote the desired function induced by the random source vector $\mathbf{S}$. In the following, we define random variables associated with the desired function, which will be used throughout the paper.

Definition 4 ($f(\mathbf{S})$ induced by $\mathbf{S}$): Let us define $U_l = \sum_{i=1}^K a_{li} S_i$ for $l \in [1 : L]$ and $B_l = \sum_{i=1}^K \mathbf{1}_{S_i=l}$ for $l \in [1 : p]$, where $L \in \mathbb{N}$. 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. ◇

B. Computing Over Fading MAC

In this subsection, we first define a block code of length n for the considered wireless INC model. Denote $\mathbf{s}_i = [s_i[1], \cdots, s_i[k]]$ and $\mathbf{x}_i = [x_i[1], \cdots, x_i[n]]$ as the length-k source vector and the length-n transmit signal vector of sensor i, respectively. Similarly, denote $\mathbf{y} = [y[1], \cdots, y[n]]$ as the length-n signal vector received at the fusion center. Then, the encoding and decoding functions of the length-n block code are given as follows.

- (Encoding) The transmit signal vector of sensor i is given by $\mathbf{x}_i = \psi(\mathbf{s}_i)$ for $i \in [1 : K]$.
- (Decoding) The fusion center estimates $\hat{f}(\mathbf{s}[j]) = \varphi_j(\mathbf{y})$ for all $j \in [1 : k]$, which implies the fusion center tries to obtain each sample-by-sample function based on the whole received signal vector of length n.

Based on the above block code, the computation rate is defined in the following.

Definition 5 (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 $\Pr\left[\bigcup_{j=1}^k \{\hat{f}(\mathbf{s}[j]) \neq f(\mathbf{s}[j])\}\right] \to 0$ as n increases.[4] ◇

From Definition 5, the computation rate R bits/channel use is the number of information bits with respect to the desired function delivered reliably by each channel use. Note that R in Definition 5 indicates the ergodic or long-term average computation rate over fading blocks. Hence, it is crucially important to design an efficient communication coding scheme in order to increase the computation rate for a given desired function over fading channels. In particular, the primary aim of this paper is to establish such a coding scheme that is able to provide a non-vanishing computation rate independent of the number of sensors K for a broad class of desired funtions by considering the function class in Definition 1.

III. Preliminaries and Main Results

In this section, we briefly introduce existing INC schemes and their limitations in fading environment, and then state our main results. We further provide some numerical results in order to demonstrate an improved computation rate from the proposed opportunistic INC. For easy presentation, we omit the time slot index t and denote the channel coefficient from sensor i by h_i in rate expressions. The statistical expectation in rate expressions is over the channel coefficients.

A. Previous Work

In [11], the computation rate for the arithmetic sum and type functions in Definition 1 has been investigated over the Gaussian (non-fading) MAC. That is the case where $h_i[t] = h_i$ in (2) and it is fixed over all time slots. The authors showed that

$$R(h_1, \cdots, h_K) = \mathsf{C}^+\left(\frac{1}{K} + \min_{i\in[1:K]} |h_i|^2 P\right) \tag{3}$$

is achievable, see [11, Theorem 3].[5] By applying [11, Theorem 3] to the fading MAC in Section II, we can show that

[3] A function is said to be symmetric if it satisfies $f(s_{\sigma(1)}, \cdots, s_{\sigma(K)}) = f(s_1, \cdots, s_K)$ for every permutation σ on $[1 : K]$.

[4] Note that $k(n)$ denotes the number of reliably computable function values with a length-n block code, which is a function of n.

[5] The computation rate in [11] is defined as $\frac{k(n)}{n}$. We present the results in [11] based on Definition 5.

$$R = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \min_{i\in[1:K]} |h_i|^2 P\right)\right] \tag{4}$$

is achievable. To improve the computation rate for fading environments, long-term power control has been considered in [11, Theorem 5], provided that

$$R = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \frac{\min_{i\in[1:K]} |h_i|^2 P}{\mathsf{E}[\min_{i\in[1:K]} |h_i|^2/|h_1|^2]}\right)\right] \tag{5}$$

is achievable, where $\frac{1}{\mathsf{E}[\min_{i\in[1:K]} |h_i|^2/|h_1|^2]} \geq 1$ is the gain from power control. Note that we assume i.i.d. fading channels over K sensors and h_1 is used as a representative coefficient without loss of generality. Another simple approach is that each sensor node transmits its source separately to the fusion center without INC, which is also known as a *time sharing* technique, achieving the computation rate of

$$R = \frac{1}{K}\mathsf{E}\left[\mathsf{C}(|h_1|^2 P)\right], \tag{6}$$

where h_1 is used again as a representative coefficient without loss of generality.

For i.i.d. Rayleigh fading channels, the computation rates in (4) to (6) achievable by the conventional INC techniques decrease as the number of sensors K increases, and eventually converge to zero for large K, see also numerical results in Section VI. To the best of our knowledge, there exists no INC framework that is able to achieve a non-vanishing computation rate for general i.i.d. fading channels in the large K regime [7], [8], [10]–[12], [20]–[22]. Therefore, the primary goal of this paper is to design an efficient INC framework that achieves a non-vanishing computation rate for general i.i.d. fading channels even if the number of sensors K in the network increases.

B. Main Results

We first present the computation rate of the proposed opportunistic INC over the fading MAC in Theorem 1 and then prove that it achieves a non-vanishing computation rate regardless of K in Corollary 1. The detailed descprition of the proposed opportunistic INC will be given in Section IV.

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

$$R = \frac{1}{N}\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 K P}{\sum_{i=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_i}|^2}\right]}\right)\right], \tag{7}$$

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

Proof: We refer to Sections IV-A to IV-C for the proof. ■

By setting $M = K$ and $N = 1$ in Theorem 1, the achievable computation rate is given by

$$R = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \frac{|h_{\pi_K}|^2 K P}{\sum_{i=1}^{K} \mathsf{E}\left[\frac{|h_{\pi_K}|^2}{|h_i|^2}\right]}\right)\right] = \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{K} + \frac{|h_{\pi_K}|^2 P}{\mathsf{E}\left[\frac{|h_{\pi_K}|^2}{|h_1|^2}\right]}\right)\right], \tag{8}$$

where the second equality follows from the fact that channel coefficients are i.i.d. Hence, Theorem 1 generalizes the result in [11, Theorem 5]. By setting $M = 1$ and $N = K$ in Theorem 1, on the other hand,

$$R = \frac{1}{K}\mathsf{E}\left[\mathsf{C}\left(|h_{\pi_1}|^2 K P\right)\right] \tag{9}$$

is achievable, which improve the computation rate in (6) by introducing opportunistic transmission and power control. For fair comparison, we will use (9) as the computation rate of the separation-based approach for the rest of this paper.

As we will explain in details in Section IV, the proposed opportunistic INC provides a general framework exploiting both *the superposition property of wireless channels*, which has been used for the INC schemes in non-fading networks [7], [8], [10], [11], and *the locally computable property of the desired function*, which has been used for INC schemes in tree networks [20], [23] and interactive computing between nodes [12], [21]. For the fading MAC, the former approach corresponds to the case when $M = 1$ and the latter approach corresponds to the case when $M = K$ in Theorem 1, respectively.

Finding the M that maximizes the computation rate in Theorem 1 is quite challenging because it depends on K and P as well as the channel distribution $f_h(\cdot)$. The following corollary shows that the proposed opportunistic INC attains a non-vanishing computation rate even if K tends to infinity by setting $M = \frac{K}{2}$. The numerical results in Section VI also demonstrate that the numerically optimal M for i.i.d. Rayleigh fading is approximately given as $\frac{K}{2}$, see Table II. The following corollary shows that M approximately given as $\frac{K}{2}$ indeed provides a non-vanishing computation rate for any i.i.d. fading channels.

Corollary 1 As K increases, the achievable computation rate in Theorem 1 converges to

$$R = \min\left(\Delta\mathsf{E}\left[\mathsf{C}\left(|h_1|^2 P\right)\right], \frac{1-\Delta}{2}\mathsf{C}^+(2\mu P)\right), \tag{10}$$

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 Section IV-D for the proof. ■

Note that Δ and μ in Corollary 1 do not depend on K, and thus R 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 equal to -1.4 dB.

In practical WSNs, on the other hand, in-network function computing over a large number of fading blocks may not be

feasible due to delay constraints, system complexity, energy efficiency, etc. In order to address such practical constraints, we also propose an opportunistic INC framework that is constrained to code only over a finite number of fading blocks and derive its computation rate under the delay constraints. The following theorem states the delay-limited computation rate of the proposed opportunistic INC.

Theorem 2 For any $M, N \in \mathbb{N}$ satisfying $MN = K$, the computation rate of the opportunistic INC over the fading MAC described in Section II under the delay constraint of N fading blocks is given by

$$R = \frac{1}{N}\mathsf{E}\left[\mathsf{C}^{+}\left(\frac{1}{M} + \frac{\min_{t\in[1:N]}\left\{\left|h_{\psi_M[t]}[t]\right|^2\right\} KP}{\sum_{t=1}^{N}\sum_{i=1}^{M}\mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_i[t]}[t]|^2}\right]}\right)\right], \tag{11}$$

where, for $t \in [1:N]$, $\{\psi_i[t]\}_{i=1}^{K-M(t-1)}$ denotes the set of sensor indices of $[1:K] \setminus \bigcup_{t'=1}^{t-1}\{\psi_i[t']\}_{i=1}^{M}$ such that $|h_{\psi_1[t]}[t]| \geq |h_{\psi_2[t]}[t]| \geq \cdots \geq |h_{\psi_{K-M(t-1)}[t]}[t]|$.[6]

Proof: We refer to Section V-A for the proof. ■

From the above definition, $\{\psi_i[t]\}_{i=1}^{K-M(t-1)}$ implies the set of sensor indices that is not included in $\bigcup_{t'=1}^{t-1}\{\psi_i[t']\}_{i=1}^{M}$, which is ordered according to their channel gains. For example, $\psi_1[t]$ indicates the sensor index having the largest channel gain among the remaining sensors after $(t-1)$ transmission durations. Note that $\mathrm{card}(\{\psi_i[t]\}_{i=1}^{K-M(t-1)})$ decreases by M for each fading block since M sensors simultaneously send their observations in each fading block.

IV. Opportunistic In-Network Computation

We first propose a novel INC framework exploiting both the opportunistic communication for fading environment and the locally computable property of the desired function in Sections IV-A to IV-C, which provides the computation rate in Theorem 1. We then analyze the computation rate of the proposed INC framework in order to prove Corollary 1 in Section IV-D.

For the conventional INC schemes explained in Section III-A, all sensor nodes simultaneously participate in the transmission, and the fusion center directly computes the desired function. For fading environments, however, the computation rates of these schemes are limited by the minimum channel gain as shown in (4) and (5), and they converge to zero as the number of sensors increases for i.i.d. Rayleigh fading. To revolve such limitation, the proposed INC technique adopts the opportunistic transmission in which only the sensors having largest channel gains participate in the transmission at each time slot. The fusion center first computes the local functions consisting of a subset of sensors, and then obtains the desired function from the local functions computed over multiple time slots.

For the proposed opportunistic INC, for M and N satisfying $MN = K$, a set of M sensors with the largest channel gains among all sensors opportunistically participate in the local function computation. Then the fusion center is able to obtain the desired function from N local functions computed at each time slot if the union of N sensor sets participating each of N local function computations is the same as the entire set of K sensors, by using the locally computable property in Definition 2. In order to compute the desired function using the locally computable property as explained above, however, the following technical issues should be addressed:

- Since a set of M sensors participating the local function computation is completely determined by the channel realization at each time slot, all possible $\binom{K}{M}$ sensor subsets, meaning that all possible $\binom{K}{M}$ local functions, will appear over large enough time slots. Hence, it should be established in detail how to pair the $\binom{K}{M}/N$ local functions in order to obtain the desired function utilizing all possible $\binom{K}{M}$ local functions.
- In order to compute sample-by-sample desired functions (see (1) at the fusion center, sample indices of each N local functions used for the desired function construction should be the same. Hence, sample indices for local computing should be carefully aligned, based on the established pairing rule for the desired function construction.
- The same philosophy of long-term power control applied in [11, Theorem 5] is also applicable to obtain similar power gain. For this case, however, long-term power control satisfying the average power constraint P is more complicated than that in [11, Theorem 5] due to opportunistic transmission.

In the following subsections, we formally address the above technical issues.

A. Opportunistic Participation and Adaptive Power Control

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 $|h_{\pi_1[t]}[t]| \geq |h_{\pi_2[t]}[t]| \geq \cdots \geq |h_{\pi_K[t]}[t]|$.

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 $P_i[t]$ denote the transmit signal power of sensor i at time slot t. We set

$$P_{\pi_i[t]}[t] = \begin{cases} c\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_i[t]}[t]|^2} & \text{for } i \in [1:M], \\ 0 & \text{otherwise}, \end{cases} \tag{12}$$

where $c > 0$ is a constant independent of channel realizations. Then

$$\begin{aligned}\mathsf{E}[P_i[t]] &= \sum_{j=1}^{K}\Pr(i=\pi_j[t])\mathsf{E}\left[P_i[t]|i=\pi_j[t]\right] \\ &\overset{(a)}{=} c\sum_{j=1}^{M}\Pr(i=\pi_j[t])\mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right] \\ &\overset{(b)}{=} \frac{c}{K}\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right],\end{aligned} \tag{13}$$

[6]For notational simplicity, we assume that $\bigcup_{i=1}^{0} A_i = \emptyset$.

where (a) follows from (12) and (b) follows since the channel coefficients are i.i.d. Hence, in order to satisfy $\mathsf{E}[P_i[t]] \le P$, we set

$$P_{\pi_i[t]}[t] = \begin{cases} \dfrac{KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]} \dfrac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_i[t]}[t]|^2} & \text{for } i \in [1:M], \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

In Section II, we assume global CSI at all sensor nodes and the fusion center for convenience, but sensor i only requires $|h_{\pi_i[t]}[t]|^2$, $|h_{\pi_M[t]}[t]|^2$, and $\{\pi_j[t]\}_{j=1}^{M}$ to implement the opportunistic power control in (14). Note that $\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]$ is a deterministic value independent of instantaneous channel gains. The following procedure briefly explains how to attain the required information.

- First, the fusion center estimates each of the K channel coefficients based on a reference signal from each sensor.
- Second, the fusion center calculates $|h_{\pi_M[t]}[t]|^2$ and $\{\pi_j[t]\}_{j=1}^{M}$ from the estimated channel coefficients and broadcasts them to all sensors.
- Third, the fusion center calculates $|h_{\pi_i[t]}[t]|^2$ from the estimated channel coefficients and sends it to sensor i for all $i \in [1:K]$.

B. Computation Rate for Local Functions

Recall that, 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 $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$. Now consider the computation rate for the above local function. Let $R'(t)$ denote the computation rate at time slot t for the local function $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$. We apply the same computing code proposed in [11] for each local function computation. We refer to [11, Theorem 3] for the detailed code construction and also refer to [11, Section IV-A] for the intuitive explanation of how to compute the arithmetic sum function over the Gaussian (nonfading) MAC. Hence, from (3), the computation rate of the local function $f(s_{\pi_1[t]}, \cdots, s_{\pi_M[t]})$ is given as

$$\begin{aligned} R'[t] &= \mathsf{C}^{+}\left(\frac{1}{M} + \min_{i\in[1:M]} |h_{\pi_i[t]}[t]|^2 P_{\pi_i[t]}[t]\right) \\ &= \mathsf{C}^{+}\left(\frac{1}{M} + \min_{i\in[1:M]} |h_{\pi_i[t]}[t]|^2 \cdot \right. \\ &\qquad \left. \frac{KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]} \frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_i[t]}[t]|^2}\right) \qquad (15) \\ &= \mathsf{C}^{+}\left(\frac{1}{M} + \frac{|h_{\pi_M[t]}[t]|^2 KP}{\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]}\right), \qquad (16) \end{aligned}$$

where the second inequality follows from (14).

C. Construction of the Desired Function Via Local Functions

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 the subsets of M sensors. Then, from the locally computable property of $f(\mathbf{s})$ in Definition 2, the fusion center is able to construct it by using $f(\{s_i\}_{i\in\lambda_1})$ to $f(\{s_i\}_{i\in\lambda_N})$ if $\bigcup_{l=1}^{N} \lambda_l = [1:K]$ is satisfied.

In order to exploit such locally computable property, however, the first two technical issues mentioned at the beginning of this section should be addressed. To address them, we first define some sensor and time index sets as follows. Let

$$\Lambda = \{\lambda \subseteq [1:K] : \mathrm{card}(\lambda) = M\} \tag{17}$$

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

$$\Omega = \left\{\omega = (\lambda_1, \cdots, \lambda_N) \in \Lambda^N : \bigcup_{l=1}^{N} \lambda_l = [1:K]\right\} \tag{18}$$

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: Before stating the proposed scheme in details, we first provide an intuitive explanation on how to handle two technical issues mentioned at the beginning of this section, i.e., about the sample-by-sample desired function computation. For easy presentation, we explain based on the case where $K = 4$ and $M = N = 2$. For this case,

$$\begin{aligned} \Lambda &= \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}, \\ \Omega &= \{((1,2), (3,4)), ((1,3), (2,4)), ((1,4), (2,3)), \\ &\qquad ((2,3), (1,4)), ((2,4), (1,3)), ((3,4), (1,2))\}. \end{aligned} \tag{19}$$

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 time slot $t = 1$, 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 source at time slot $t = 1$ since $(1, 3)$ firstly appears in the second element in Ω. Similarly, at time slot $t = 2$, 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 time

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

slot $t=10$, 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: For notational convenience, $f(\{\mathbf{s}_i\}_{i\in\lambda})$ denotes the set of sample-by-sample local functions, i.e.,

$$f(\{\mathbf{s}_i\}_{i\in\lambda}) = \{f(\{s_i[1]\}_{i\in\lambda}), \cdots, f(\{s_i[k]\}_{i\in\lambda})\}, \tag{20}$$

where $\lambda \in \Lambda$. 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

$$\left| \frac{1}{n}\mathrm{card}(\mathcal{T}_\lambda) - \frac{1}{\binom{K}{M}} \right| \geq \epsilon \tag{21}$$

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

$$\Pr(\{\pi_i[t]\}_{i=1}^M = \lambda) = \frac{1}{\binom{K}{M}} \tag{22}$$

for all $\lambda \in \Lambda$. Therefore, from the strong typicality in [24, 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

$$T := \frac{\frac{n}{\binom{K}{M}} - \frac{n}{\log n}}{\mathrm{card}(\Omega_\lambda)} = \frac{\frac{n}{\binom{K}{M}} - \frac{n}{\log n}}{N\prod_{l=1}^{N-1}\binom{K-Ml}{M}} \tag{23}$$

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 $\omega_1 \neq \omega_2 \in \Omega_\lambda$.

Suppose that ω is the lth element in Ω. Then, for $i \in [1:K]$, define the length-U source vector $\mathbf{s}_i(\omega) = [s_i[(l-1)U+1], \cdots, s_i[lU]]$ consisting of the $((l-1)U+1)$th source to the (lU)th source of sensor i, where $U \in \mathbb{N}$ will be determined later.

For each $\omega \in \Omega$, the time slots in $\mathcal{T}_{\lambda,\omega}$ are used to compute the set of sample-by-sample local functions $f(\{\mathbf{s}_i(\omega)\}_{i\in\lambda})$ for all $\lambda \in \omega$, see the definition in (20). Then the sample-by-sample desired functions $f(\{\mathbf{s}_i(\omega)\}_{i\in[1:K]})$ can be attained based on the computed sample-by-sample local functions.

Let us first explain how to compute $f(\{\mathbf{s}_i(\omega)\}_{i\in\lambda})$ using the time slots in $\mathcal{T}_{\lambda,\omega}$ for given $\omega \in \Omega$ and $\lambda \in \omega$. Let $\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega)) \in \mathbb{C}^{T\times 1}$ denote the length-T time-extended transmit signal vector of sensor i during $t \in \mathcal{T}_{\lambda,\omega}$, where $i \in [1:K]$. That is, the set of sources $\mathbf{s}_i(\omega)$ is mapped into $\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega))$ and transmitted during the time slots in $\mathcal{T}_{\lambda,\omega}$. Recall that, at each time slot $t \in \mathcal{T}_{\lambda,\omega}$, the transmit power of each element in $\mathbf{x}_i(\lambda,\omega)$ is adjusted based on (14). Specifically, we construct

$$\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega)) = \Gamma_{i,\lambda,\omega}\mathbf{x}'_{i,\lambda,\omega}(\mathbf{s}_i(\omega)), \tag{24}$$

where $\Gamma_{i,\lambda,\omega} = \mathrm{diag}\left(\left\{\sqrt{P_i[t]}\frac{|h_i[t]|}{h_i[t]}\right\}_{t\in\mathcal{T}_{\lambda,\omega}}\right)$ and $\mathbf{x}'_{i,\lambda,\omega}(\mathbf{s}_i(\omega))$ is the lattice-based transmit signal vector for the distributed INC satisfying the average transmit power of one used in [11, Theorem 3], see also [8], [13], [14] for the lattice construction. Here, the definition of $P_i[t]$ is given by

$$P_i[t] = \begin{cases} \frac{KP}{\sum_{j=1}^M \mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]}\frac{|h_{\pi_M[t]}[t]|^2}{|h_i[t]|^2} & \text{for } i \in \{\pi_j[t]\}_{j=1}^M, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

from (14). Obviously, $\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega)) = \mathbf{0}$ from (25) if $i \notin \lambda$ since the M sensors with the largest channel gains are in λ for $t \in \mathcal{T}_{\lambda,\omega}$. Then, sensor i transmits $\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega))$ during $t \in \mathcal{T}_{\lambda,\omega}$ for all $i \in [1:K]$.

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

$$\mathbf{y}_{\lambda,\omega} = \sum_{i=1}^K \mathbf{H}_{i,\lambda,\omega}\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega)) + \mathbf{z}_{\lambda,\omega} = \sum_{i\in\lambda} \mathbf{H}_{i,\lambda,\omega}\mathbf{x}_{i,\lambda,\omega}(\mathbf{s}_i(\omega)) + \mathbf{z}_{\lambda,\omega}, \tag{26}$$

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 (24),

$$\begin{aligned}\mathbf{y}_{\lambda,\omega} &= \sum_{i\in\lambda} \mathbf{H}_{i,\lambda,\omega}\Gamma_{i,\lambda,\omega}\mathbf{x}'_{i,\lambda,\omega} + \mathbf{z}_{\lambda,\omega} \\ &= \sum_{i\in\lambda} \mathbf{H}'_{i,\lambda,\omega}\mathbf{x}'_{i,\lambda,\omega} + \mathbf{z}_{\lambda,\omega}, \end{aligned} \tag{27}$$

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

$$\begin{aligned} h'_i[t] &= h_i[t]\sqrt{P_i[t]}\frac{|h_i[t]|}{h_i[t]} \\ &= \sqrt{\frac{KP}{\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\pi_M[t]}[t]|^2}{|h_{\pi_j[t]}[t]|^2}\right]}}|h_{\pi_M[t]}[t]|. \end{aligned} \tag{28}$$

Therefore, from (15), the computation rate of

$$\frac{1}{T}\sum_{t\in\mathcal{T}_{\lambda,\omega}} \mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}[t]|^2 KP}{\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\pi_M}[t]|^2}{|h_{\pi_j}[t]|^2}\right]}\right) \tag{29}$$

is achievable for computing the local function $f(\{s_i\}_{i\in\lambda})$, which converges to the ergodic rate of

$$R' := \mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 KP}{\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_j}|^2}\right]}\right)\right] \tag{30}$$

as n increase because $\mathrm{card}(\mathcal{T}_{\lambda,\omega}) = T$ increases as n increases. More specifically, the fusion center is able to compute $f(\{\mathbf{s}_i(\omega)\}_{i\in\lambda})$ by setting $U = \frac{R'T}{H(f(\mathbf{S}))}$ during $t \in \mathcal{T}_{\lambda,\omega}$.

Let us now explain how to construct $f(\{\mathbf{s}_i(\omega)\}_{i\in[1:K]})$ using the time slots in $\bigcup_{\lambda\in\Lambda:\lambda\in\omega}\mathcal{T}_{\lambda,\omega}$ for given $\omega \in \Omega$. From now on, assume $U = \frac{R'T}{H(f(\mathbf{S}))}$. During $t \in \bigcup_{\lambda\in\Lambda:\lambda\in\omega}\mathcal{T}_{\lambda,\omega}$, the fusion center first computes $\{f(\{\mathbf{s}_i(\omega)\}_{i\in\lambda})\}_{\lambda\in\Lambda:\lambda\in\omega}$. Notice that, from the definition of $\mathbf{s}_i(\omega)$, the source indices are aligned for the same ω. Hence, by using the locally computable property of the desired function in Definition 2, the fusion center is able to attain $f(\{\mathbf{s}_i(\omega)\}_{i\in[1:K]})$ from $\{f(\{\mathbf{s}_i(\omega)\}_{i\in\lambda})\}_{\lambda\in\Lambda:\lambda\in\omega}$ during $t \in \bigcup_{\lambda\in\Lambda:\lambda\in\omega}\mathcal{T}_{\lambda,\omega}$.

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

$$\begin{aligned} k(n) &= \left(\prod_{l=0}^{N-1}\binom{K-Ml}{M}\right)U \\ &= \left(\prod_{l=0}^{N-1}\binom{K-Ml}{M}\right)\frac{R'T}{H(f(\mathbf{S}))} \end{aligned} \tag{31}$$

and, as a result, the computation rate of

$$\begin{aligned} R &= \frac{k(n)H(f(\mathbf{S}))}{n} \\ &\overset{(a)}{=} \frac{\left(\prod_{l=0}^{N-1}\binom{K-Ml}{M}\right)R'T}{n} \\ &\overset{(b)}{=} \frac{\left(\prod_{l=0}^{N-1}\binom{K-Ml}{M}\right)R'}{n}\frac{\frac{n}{\binom{K}{M}} - \frac{n}{\log n}}{N\prod_{l=1}^{N-1}\binom{K-Ml}{M}} \\ &= \frac{1}{N}R'\left(1 - \frac{\binom{K}{M}}{\log n}\right) \end{aligned} \tag{32}$$

is achievable, where (a) and (b) follow from (31) and (23) respectively. In conclusion, (7) is achievable as n increases, which complete the proof of Theorem 1.

D. Non-Vanishing Computation Rate

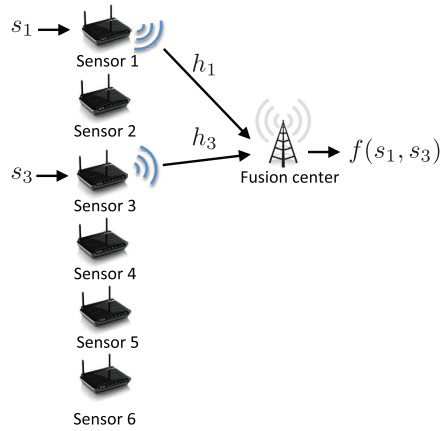
In this subsection, we prove Corollary 1. Suppose that K is even. Let $\bar{\mu}_K$ denote the sample median for $\{|h_i|^2\}_{i\in[1:K]}$. By setting $M = \frac{K}{2}$ and $N = 2$ in Theorem 1, we have

$$\begin{aligned} R &= \frac{1}{2}\mathsf{E}\left[\mathsf{C}^+\left(\frac{2}{K} + \frac{|h_{\pi_{K/2}}|^2 KP}{\sum_{i=1}^{K/2}\mathsf{E}\left[\frac{|h_{\pi_{K/2}}|^2}{|h_{\pi_i}|^2}\right]}\right)\right] \\ &\overset{(a)}{\geq} \frac{1}{2}\mathsf{E}\left[\mathsf{C}^+\left(\frac{2}{K} + 2|h_{\pi_{K/2}}|^2 P\right)\right] \\ &\overset{(b)}{\geq} \frac{1}{2}\mathsf{E}\left[\mathsf{C}^+\left(\frac{2}{K} + 2\bar{\mu}_K P\right)\right], \end{aligned} \tag{33}$$

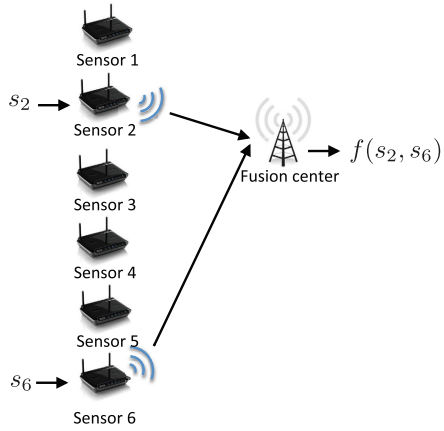
where (a) follows since $|h_{\pi_i}|^2 \geq |h_{\pi_{K/2}}|^2$ for all $i \in [1 : K/2]$ and (b) follows since $\bar{\mu}_K = \frac{1}{2}(|h_{\pi_{K/2}}|^2 + |h_{\pi_{K/2-1}}|^2)$ for even K. As K increases, $\bar{\mu}_K$ asymptotically follows the Gaussian distribution with mean μ and variance $\frac{1}{4K(f_{|h|^2}(\mu))^2}$ [25]. Hence, from (33), we have $\lim_{K\to\infty} R = \frac{1}{2}\mathsf{C}^+(2\mu P)$ for even K. Now suppose that K is odd. Then the first sensor transmits its source to the fusion center using a Δ fraction of time slots, which at least achieves a rate of $\Delta\mathsf{E}\left[\mathsf{C}\left(|h_1|^2 P\right)\right]$. For the rest of the sensors, the fusion center computes $f(s_2, s_3, \cdots, s_K)$ using the rest of $1-\Delta$ fraction of time slots, which at least achieves a rate of $\frac{1-\Delta}{2}\mathsf{E}\left[\mathsf{C}^+\left(\frac{2}{K-1} + 2\bar{\mu}_{K-1}P\right)\right]$ from (33). Then finally, the fusion center computes the desired function using the locally computable property. i.e., $f(\mathbf{s}) = g(s_1, f(s_2, \cdots, s_K))$. Hence

$$\begin{aligned} R = \min\Bigg(&\Delta\mathsf{E}\left[\mathsf{C}\left(|h_1|^2 P\right)\right], \\ &\frac{1-\Delta}{2}\mathsf{E}\left[\mathsf{C}^+\left(\frac{2}{K-1} + 2\bar{\mu}_{K-1}P\right)\right]\Bigg) \end{aligned} \tag{34}$$

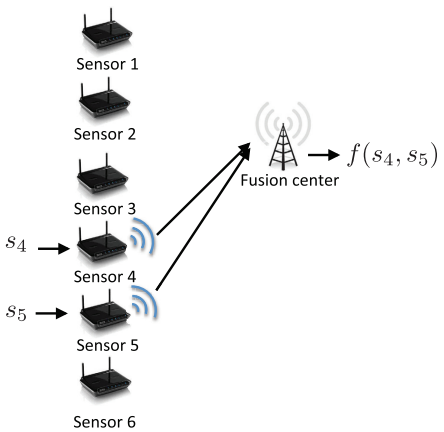
is achievable for odd K, which converges to $\min\left(\Delta\mathsf{E}\left[\mathsf{C}\left(|h_1|^2 P\right)\right], \frac{1-\Delta}{2}\mathsf{C}^+(2\mu P)\right)$ as K increases. In conclusion, Corollary 1 holds.



(a) Local function computing at the first time slot when $|h_1|, |h_3| \geq |h_2|, |h_4|, |h_5|, |h_6|$



(b) Local function computing at the seccond time slot when $|h_2|, |h_6| \geq |h_4|, |h_5|$



(c) Local function computing at the thrid time slot

Fig. 2. Opportunistic wireless INC under delay constraint.

V. Opportunistic In-Network Computation Under Delay Constraint

In this section, we consider opportunistic INC under delay constraint, which is defined as the number of required fading blocks for INC. Counting the number of required fading blocks is meaningful in practical systems because it is related not only to the delay normalized by coding or fading blocks but also overall implementation complexity [26]–[28].

In order to address such an delay issue, we propose a modified version of the opportunistic INC depicted in Fig. 2. For example, assume that $K = 6$, $M = 2$, and $N = 3$. As the same manner in Section IV, the M sensors having the largest channel gains participate in the local function computation at the first time slot. Suppose that sensors 1 and 3 participate and the fusion center computes $f(s_1, s_3)$ at the first time slot. Then at the second time slot, the sensors except 1 and 3 opportunistically participate and suppose that sensors 2 and 6 participate and the fusion center computes $f(s_2, s_6)$ at the second time slot. Finally, at the third time slot, sensors 4 and 5 participate and the fusion center computes $f(s_4, s_5)$. Then, the fusion center attains $f(s_1, \cdots, s_6)$ from the computed $f(s_1, s_3)$, $f(s_2, s_6)$, and $f(s_4, s_5)$ from the locally computable property (Definition 2), which guarantees the delay of N time slots.

A. Proof of Theorem 2

In this subsection, we prove Theorem 2. At each time slot $t \in [1 : N]$, the M sensors in $\{\psi_i[t]\}_{i=1}^M$ participate in the transmission and the fusion center computes the local function $f(s_{\psi_1[t]}, s_{\psi_2[t]}, \cdots, s_{\psi_M[t]})$, see the definition of $\{\psi_i[t]\}_{i=1}^M$ in Theorem 2. That is, at each time slot $t \in [1 : N]$, the M sensors with the largest channel gains among the sensors that did not participate in the previous slots participate in the transmission. Let $P_i[t]$ denote the transmit signal power of sensor i at time slot t. We set

$$P_{\psi_i[t]}[t] = \begin{cases} c\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_i[t]}[t]|^2} & \text{for } i \in [1 : M], \\ 0 & \text{otherwise,} \end{cases} \tag{35}$$

where $c > 0$ is a constant independent of channel realizations. Then

$$\begin{aligned} \mathsf{E}[P_i[t]] &= \sum_{t=1}^{N}\sum_{j=1}^{M} \Pr(i = \psi_j[t]) \mathsf{E}\left[P_i[t] | i = \psi_j[t]\right] \\ &\overset{(a)}{=} c\sum_{t=1}^{N}\sum_{j=1}^{M} \Pr(i = \psi_j[t]) \mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_j[t]}[t]|^2}\right] \end{aligned} \tag{36}$$

where the second equality follows from (35). For $t = 1$, $\Pr(i = \psi_j[t]) = \frac{1}{K}$. For $t \in [2 : N]$, we have

$$\begin{aligned} \Pr(i = \psi_j[t]) = &\left(\prod_{k=1}^{t-1} \Pr\left(i \notin \{\psi_j[k]\}_{j=1}^M | i \notin \{\psi_j[k-1]\}_{j=1}^M, \right.\right. \\ &\left.\left. \cdots, i \notin \{\psi_j[1]\}_{j=1}^M\right)\right) \\ &\cdot \Pr(i = \psi_j[t] | i \notin \{\psi_j[t-1]\}_{j=1}^M, \cdots, i \notin \{\psi_j[1]\}_{j=1}^M) \\ = &\left(\prod_{k=1}^{t-1} \frac{K - Mk}{K - M(k-1)}\right) \frac{1}{K - M(t-1)} = \frac{1}{K} \end{aligned} \tag{37}$$

Hence, $\Pr(i = \psi_j[t]) = \frac{1}{K}$ for $t \in [1 : N]$ and then from (36)

$$\mathsf{E}[P_i[t]] = \frac{c}{K}\sum_{t=1}^{N}\sum_{j=1}^{M} \mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_j[t]}[t]|^2}\right]. \tag{38}$$

Hence in order to satisfy $\mathsf{E}[P_i[t]] \le P$, we set $P_{\psi_i[t]}[t]$ as

$$\begin{cases} \dfrac{KP}{\sum_{t=1}^{N}\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_j[t]}[t]|^2}\right]}\dfrac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_i[t]}[t]|^2} & \text{for } i \in [1:M], \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

Now consider the computation rate for the local function at time slot t, i.e., $f(\{s_{\psi_i[t]}\}_{i=1}^{M})$. In a similar manner as in (15), from (39), the computation rate of

$$\begin{aligned} R''[t] &= \mathsf{C}^+\left(\frac{1}{M} + \min_{i\in[1:M]} |h_{\psi_i[t]}[t]|^2 P_{\psi_i[t]}[t]\right) \\ &= \mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\psi_M[t]}[t]|^2 KP}{\sum_{t=1}^{N}\sum_{j=1}^{M}\mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_j[t]}[t]|^2}\right]}\right). \end{aligned} \tag{40}$$

is achievable for computing $f(\{s_{\psi_i[t]}\}_{i=1}^{M})$. Recall that $\bigcup_{t=1}^{N}\{\psi_i[t]\}_{i=1}^{M} = [1:K]$, which means that all sensors participate in the local function computation once during N time slots. Hence the fusion center is able to attain the desired function from the locally computable property, i.e., $f(\mathbf{s}) = g(f(\{s_{\psi_i[1]}\}_{i=1}^{M}), \cdots, f(\{s_{\psi_i[N]}\}_{i=1}^{M}))$. That is, the computation rate of

$$\begin{aligned} R &= \frac{1}{N}\min_{t\in[1:N]}\left\{R''[t]\right\} \\ &= \frac{1}{N}\mathsf{C}^+\left(\frac{1}{M} + \frac{\min_{t\in[1:N]}\left\{\left|h_{\psi_M[t]}[t]\right|^2\right\} KP}{\sum_{t=1}^{N}\sum_{i=1}^{M}\mathsf{E}\left[\frac{|h_{\psi_M[t]}[t]|^2}{|h_{\psi_i[t]}[t]|^2}\right]}\right) \end{aligned} \tag{41}$$

is achievable within the delay of N time slots for computing each sample-by-sample desired function. By applying the above opportunistic INC scheme repeatedly over large enough blocks, the computation rate of (11) is achievable within the delay of N time slots (fading blocks) for computing each sample-by-sample desired function. In conclusion, Theorem 2 is proved.

VI. Numerical Results

In this section, we provide several numerical examples of the computation rates of the opportunistic INC and compare them with the computation rates of the conventional INC schemes.

Fig. 3 shows the computation rate in Theorem 1 with respect to M. The results demonstrate that the computation rate in Theorem 1 with optimally chosen M outperforms the conventional approaches, which are the cases where $M = 1$ and $M = K$ in the figure. In addition, the optimal M resulting in the largest computation rate varies according to the average transmit power P.[7] The computation rate certainly increases as P increases.

Fig. 4 shows the computation rate in Theorem 1 with optimal M as the number of sensors in the network increases. As shown in the figure, the proposed INC with optimal M achieves a non-vanishing computation rate even if K increases,

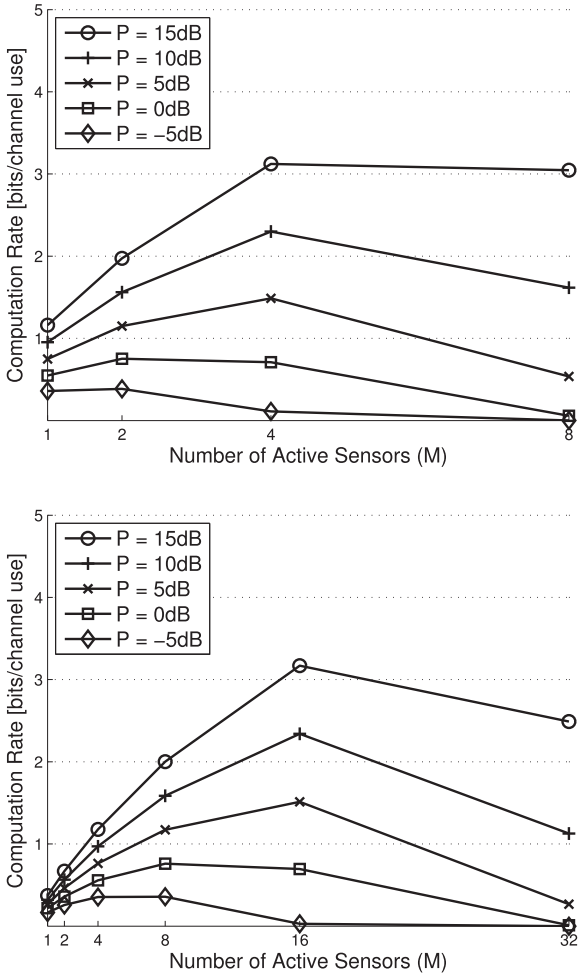


Fig. 3. Computation rates with respect to M when $K = 8$ (top) and $K = 32$ (bottom) for i.i.d. Rayleigh fading.

while the computation rates attained by the conventional INC schemes (the cases where $M = 1$ and $M = K$) converge to zero as K increases. The formal proof of achieving a non-vanishing computation rate is given in Corollary 1. Notice that this improvement comes from the opportunistic scheduling gain and the optimal power adjustment. Therefore, the computation rate gap between the proposed INC and the conventional INC schemes becomes significant as the number of sensors K increases. The numerically optimal M maximizing the computation rate in Theorem 1 is given in Table II, which is a function of K and P. As seen in Table II, the optimal M increases as P increases and, more importantly, it scales approximately as $\frac{K}{2}$ as K increases. Table III shows the average required transmit power P at each sensor for achieving given computation rates when $K = 16$ and $K = 64$. As seen in the table, the proposed opportunistic INC can significantly save the average transmit power compared with the conventional INC schemes. For instance, it achieves $R = 1$ bits/sec/Hz with 22% and 18% of power consumption compared with the conventional INC scheme in [11] (the case where $M = K$).

Fig. 5 compares the computation rates of the proposed INC with/without delay contraints in Theorems 1 and 2 with respect to the average transmit power P. For both cases, the number of active sensors in each slot M is optimally chosen. Hence,

[7]The average signal-to-noise ratio (SNR) is the same as P since the noise variance is given by one in Section II.

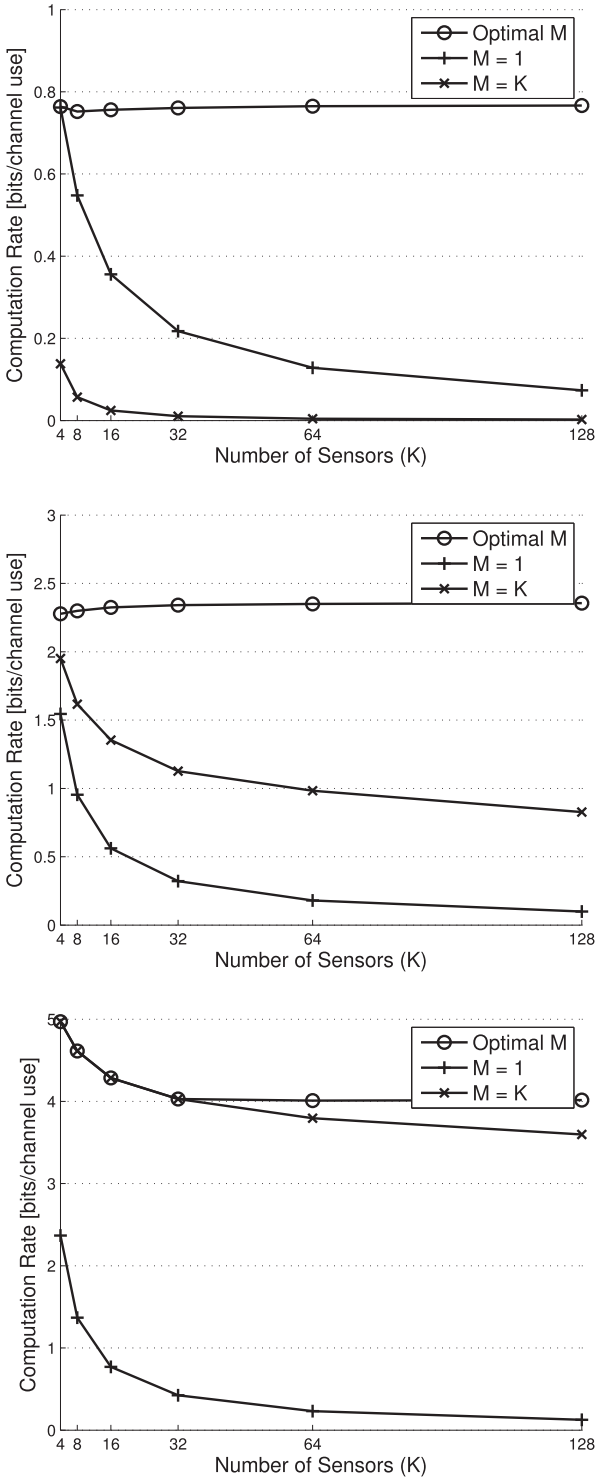


Fig. 4. Computation rates with respect to K when $P = 0$ dB (top), $P = 10$ dB (middle), and $P = 20$ dB (bottom) for i.i.d. Rayleigh fading.

Theorem 2 is able to compute each sample-by-sample desired function within the delay of at most K fading blocks since $N \leq K$. Although Theorem 1 provides a better computation rate than Theorem 2, the delay for computing each sample-by-sample desired function can be arbitrarily large. As seen in the figure, Theorem 2 guarantees a fixed and finite delay (less than K fading blocks) with the cost of small computation rate degradation.

We further compare the computation rates of the proposed INC with/without delay constraint with the conventional INC schemes in Fig. 6. Note that the INC scheme in [11, Theorem 5] computes the desired function using a single fading block, i.e., delay of one fading block, which corresponds to the case where $M = K$ in Theorem 1 (or the case where $M = K$ in Theorem 2), see also the equivalent rate expression in (5). As seen in the figure, the proposed INC scheme in Theorem 2 outperforms the previous approaches ($M = K$ and $M = 1$ with and without delay constraint), providing the guaranteed delay of at most K fading blocks. In addition, the computation rate gap from Theorem 1 with optimally chosen M decreases as K increases.

Fig. 7 compares the proposed opportunistic INC schemes with the non-opportunistic INC scheme, i.e., M sensors are chosen uniformly at random at each time slot and the fusion center performs the local function computation of the chosen M sources. Then with the same construction of the desired function via local functions in Section IV-C, the computation rate of

$$R = \frac{1}{N}\mathsf{E}\left[\mathsf{C}^{+}\left(\frac{1}{M} + \frac{\min_{i\in[1:M]}|h_i|^2 NP}{\mathsf{E}\left[\frac{\min_{i\in[1:M]}|h_i|^2}{|h_1|^2}\right]}\right)\right] \tag{42}$$

is achievable for any $M, N \in \mathbb{N}$ satisfying $MN = K$ from Theorem 1. Notice that we can also derive (42) from Theorem 2 by randomly selecting M sensors that did not participate in the previous slots. In the figure, for each K, we set M to numerically maximize the computation rate of each scheme. As seen in the figure, the computation rate of the non-opportunistic INC scheme decreases as K increases. Therefore, opportunistic sensor selection based on channel gains is essentially required to achieve a non-vanishing computation rate with increasing K.

VII. Discussions

A. Scaling Laws

It has been shown in [11, Theorem 6] that the computation rate in (5), established in [11, Theorem 5], achieves the computation capacity within a constant bps/Hz gap regardless of P for i.i.d. Rayleigh fading, which provides a universal performance guarantee for any P when the number of sensors K is fixed and finite.

Let us now focus on the asymptotic behaviour of the computation capacity as K increases when P is fixed and finite. As the same manner in [11, Section III-D], by allowing full cooperation between K sensors, the computation capacity is upper bounded by

$$\mathsf{E}\left[\log\left(1 + \left(\sum_{i=1}^{K}|h_i[t]|\phi_i^*[t]\right)^2\right)\right] \tag{43}$$

$$\overset{(a)}{\leq} 2\mathsf{E}\left[\log\left(1 + \sum_{i=1}^{K}|h_i[t]|\phi_i^*[t]\right)\right]$$

$$\overset{(b)}{\leq} 2\log\left(1 + \sum_{i=1}^{K}\mathsf{E}\left[|h_i[t]|\phi_i^*[t]\right]\right)$$

$$\overset{(c)}{\leq} 2\log\left(1 + KP\right), \tag{44}$$

TABLE II
OPTIMAL M MAXIMIZING THE COMPUTATION RATE IN THEOREM 1 FOR I.I.D. RAYLEIGH FADING

	$K=4$	$K=8$	$K=16$	$K=32$	$K=64$	$K=128$	$K=256$	$K=512$
$P=0$ dB	2	2	4	8	16	32	64	128
$P=10$ dB	2	4	8	16	32	64	128	256
$P=20$ dB	4	8	16	32	32	64	128	256

TABLE III
REQUIRED AVERAGE TRANSMIT POWER P AT EACH SENSOR FOR ACHIEVING A GIVEN COMPUTATION RATE

	$K=16$			$K=64$		
	Optimal M	$M=1$	$M=K$	Optimal M	$M=1$	$M=K$
$R=0.5$ bits/sec/Hz	-3.2 dB	6.2 dB	5.8 dB	-3.2 dB	73 dB	7.8 dB
$R=1$ bits/sec/Hz	1.9 dB	31 dB	8.4 dB	-1.9 dB	$\gg$ 100 dB	10 dB
$R=2$ bits/sec/Hz	8.0 dB	79 dB	12 dB	8.0 dB	$\gg$ 100 dB	14 dB

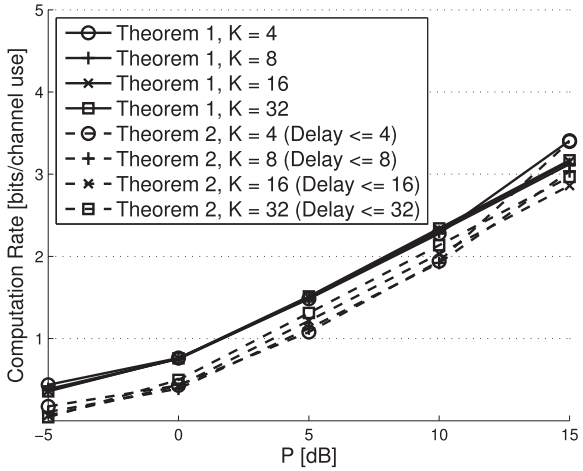


Fig. 5. Computation rates with and without delay constraint for i.i.d. Rayleigh fading.

where $\phi_i^*[t]$ denotes the optimal power allocation of sensor i, which is a function of $\{h_i[t]\}_{i=1}^K$. Here (a) follows since $\log(a+b) \le 2\log(\sqrt{a}+\sqrt{b})$ for any $a, b \ge 0$, (b) follows from Jensen's inequality, and (c) follows since $\mathsf{E}[|h_i[t]|\phi_i^*[t]] \le \sqrt{\mathsf{E}[|h_i[t]|^2]\mathsf{E}[(\phi_i^*[t])^2]} \le P$. Therefore, the computation capacity scales as at most $\Theta(\log(K))$ as K increases. However, it is quite challenging to prove whether the computation rate in (5) provides $\Theta(1)$ scaling law or not when $\min_{i\in[1:K]} |h_i[t]|^2$ converges to zero as K increases.

In Theorem 1 and Corollary 1, we have shown that the computation rate of the proposed opportunistic INC scales as $\Theta(1)$ as the number of sensors K increases for any i.i.d. fading environment if P is larger than some constant independent of P. Therefore, our result characterizes that the computation capacity is given by $\Omega(1)$ and $O(\log K)$, which tighten an achievable lower bound as $\Omega(1)$ by proposing opportunistic INC.

B. Asymmetric Channels

In order to verify the benefit of opportunistic INC for general asymmetric channels, we consider the geometric network model in this subsection. Specifically, assuming that the distance between sensor i and the fusion center is given by d_i, the received signal of the fusion center at the tth time slot is represented as

$$y[t] = \sum_{i=1}^{K} \frac{h_i[t]}{d_i^{\gamma/2}} x_i[t] + z[t], \tag{45}$$

where $\gamma \ge 2$ denotes the path-loss exponent and $h_i[t]$ denotes block fading assumed as the same manner in (2). Therefore, the same opportunistic INC in Section IV is applicable for asymmetric channels based on $\{h_i[t]\}_{i=1}^K$ at each time slot, i.e., the opportunism based on the relative received signal strength normalized by the average received signal strength at each sensor. Then, from Theorem 1, the achievable computation rate is straightforwardly given as

$$R = \frac{1}{N}\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\pi_M}|^2 K(P/d_{\max}^{\gamma})}{\sum_{i=1}^{M}\mathsf{E}\left[\frac{|h_{\pi_M}|^2}{|h_{\pi_i}|^2}\right]}\right)\right] \tag{46}$$

for any $M, N \in \mathbb{N}$ satisfying $MN = K$, where $d_{\max} = \max_{i\in[1:K]}\{d_i\}$ and the definition of $\{\pi_i\}_{i=1}^K$ is given in Theorem 1. That is, sensor i transmits with a reduced average power of $P(\frac{d_i}{d_{\max}})^{\gamma}$ because the computation rate is restricted by the sensor located at the maximum distance from the fusion center.

In simulation, we focus on two random geometric networks: a set of K sensors are uniformly deployed at random in a 1×1 square area for the first model and in a $\sqrt{K} \times \sqrt{K}$ square area for the second model, similar to dense and extended networks respectively in [29]–[31]. For both models, the fusion center is assumed to be located at the network center.

Fig. 8 plots the computation rate in (46) with optimal M as the number of sensors K in the network increases for both models. For comparison, we also plot the the conventional INC schemes (the cases where $M=1$ and $M=K$). In simulation, we plot the computation rates averaged out over large enough network realizations. Similar to Fig. 4, the proposed INC outperforms the conventional approaches due to the opportunistic gain for function computing. In particular, for the first model, the distance of each sensor from the fusion center is upper bounded by a some constant independent of K, which provides a non-vanishing computation rate even for asymmetric configurations. For the second model, on the other hand,

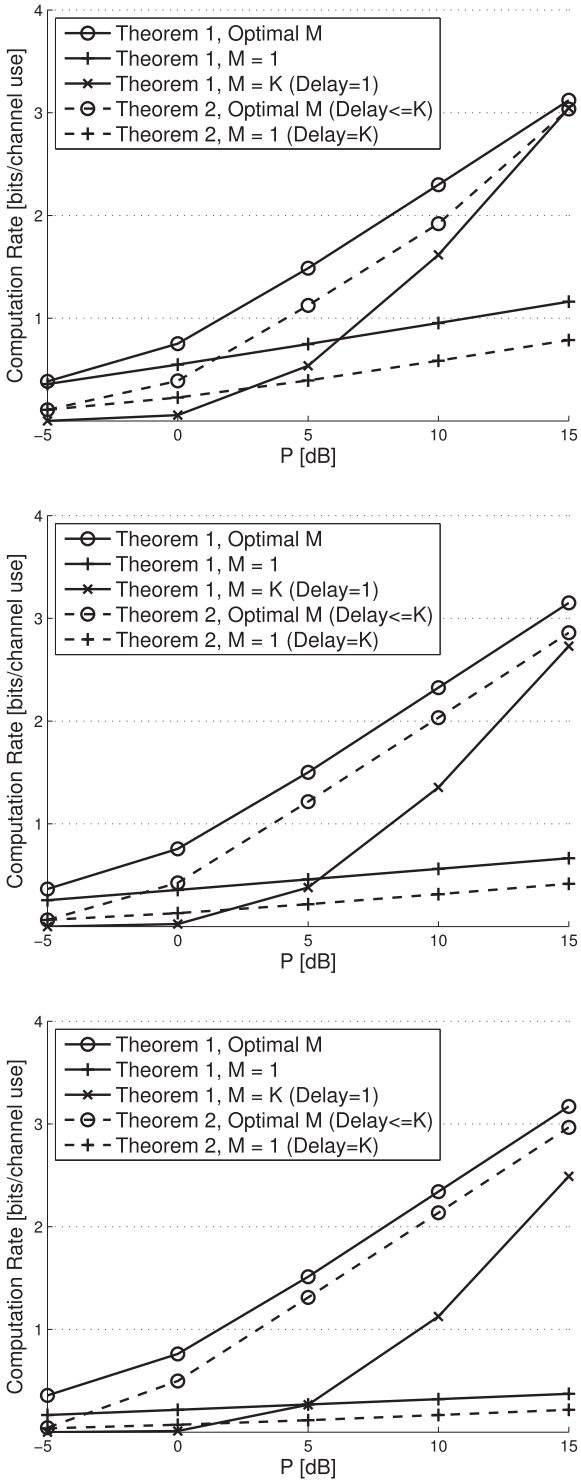


Fig. 6. Computation rates with and without delay constraint when $K = 8$ (top), $K = 16$ (middle), and $K = 32$ (bottom) for i.i.d. Rayleigh fading.

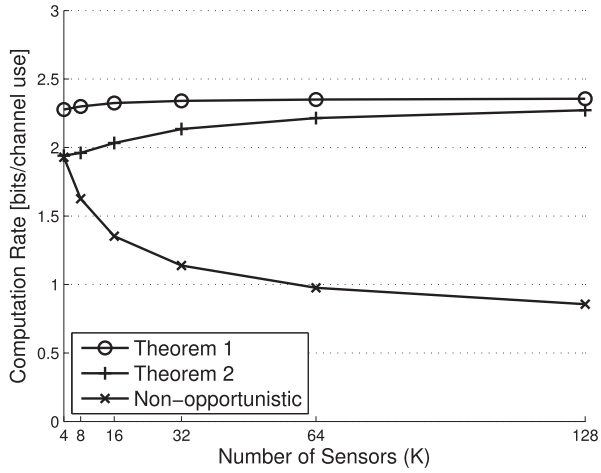


Fig. 7. Computation rates of the non-opportunistic INC with respect to K when $P = 10$ dB for i.i.d. Rayleigh fading.

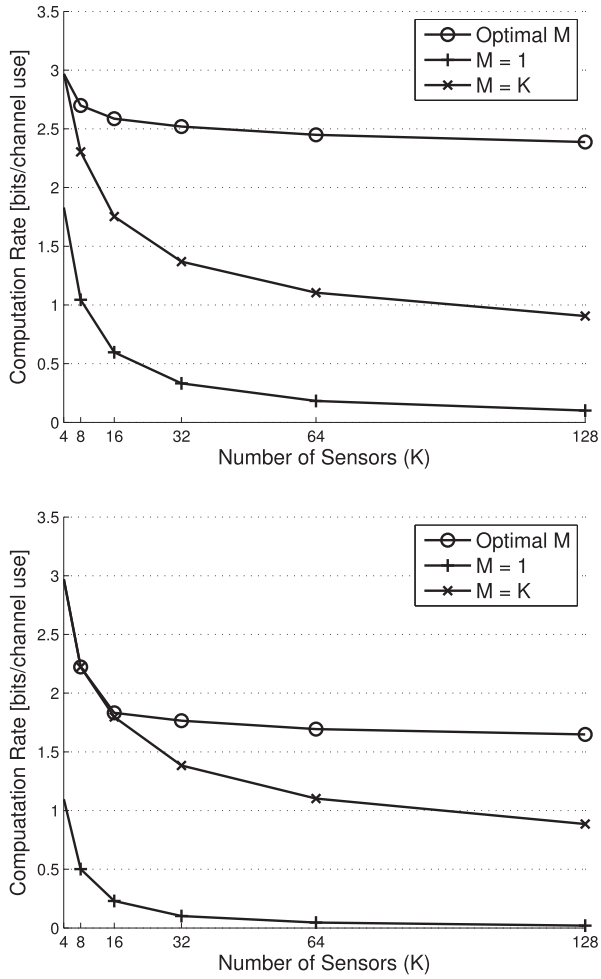


Fig. 8. Computation rates with respect to K for the first model (top) and the second model (bottom) when $P = 5$ dB and $\gamma = 3$ for i.i.d. Rayleigh fading.

the maximum distance from the fusion center increases as K increases, which let the computation rate eventually converge to zero as K increases. In such case, if sensors can communicate to each other, then the computation rate might be improved by combining opportunistic INC with multihop routing, which will be discussed in the next subsection.

C. Multihop Networks

In many practical cases of interest, some sensors might not have direct communication links to the fusion center or they might not want to directly communicate with the fusion center

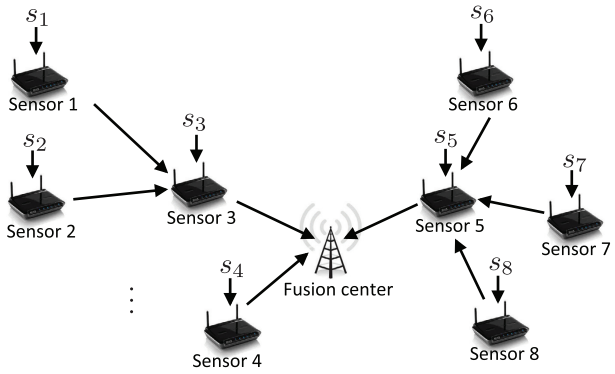


Fig. 9. An example sensor network consisting of 8 sensors and a fusion center.

due to severe signal attenuation by long-distance communication. In such cases, multihop routing using other sensors as relays can improve the computation rate. In this subsection, we briefly explain how to combine the proposed INC with a general multihop routing framework. Fig. 9 provides a high-level illustration based on an example network consisting of 8 sensors and a fusion center. For instance, sensors 1 and 2 are able to communicate with sensor 3 via the fading MAC as in Section II-A. First, sensor 3 computes the local function $f(s_1, s_2)$ by opportunistic INC from sensors 1 and 2 and, similarly, sensor 5 computes the local function $f(s_6, \cdots, s_8)$ by opportunistic INC from sensors 6, 7, and 8. Then sensors 3 and 5 can construct $f(s_1, \cdots, s_3)$ and $f(s_5, \cdots, s_8)$ respectively from their own sources. Lastly, the fusion center can attain the desired function $f(s_1, \cdots, s_8)$ again by opportunistic INC from sensors 3, 4, and 5. This is possible because each sensor first constructs the corresponding local function and compress it in the proposed INC, see [11, Theorem 3] for the detailed code construction. That is, even if sensor 3 attains s_1, s_2, s_3, it first constructs $f(s_1, \cdots, s_3)$ to apply the proposed INC. As studied in [32], [33] and the references therein, appropriate scheduling and power control can construct spanning trees for general sensor networks. Then, we can apply the above INC scheme on the top of constructed tree networks.

VIII. Concluding Remarks

In this paper, we investigated the function computation problem in WSNs, focusing on the design of an efficient INC strategy for fading environment. The proposed INC framework exploits both the superposition property of wireless channel and the locally computable property of the desired function, combined with opportunistic transmission. We showed that a non-vanishing computation rate is achievable by the opportunsitic INC even if the number of sensors in the network increases, which is the first theoretic result. Although we considered the fading MAC, the proposed INC is straightforwardly applicable to wireless networks having orthogonal components and tree networks, attaining improved computation rates for a broad class of WSNs.

References

[1] A. Giridhar and P. R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, no. 4, pp. 98–107, Apr. 2006.
[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.
[3] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 219–221, Mar. 1979.
[4] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 396–412, May 1983.
[5] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.
[6] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5247–5251, Nov. 2008.
[7] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
[8] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.
[9] R. Soundararajan and S. Vishwanath, "Communicating linear functions of correlated Gaussian sources over a MAC," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1853–1860, Mar. 2012.
[10] J. Zhan, S. Y. Park, M. Gastpar, and A. Sahai, "Linear function computation in networks: Duality and constant gap results," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 620–638, Apr. 2013.
[11] S.-W. Jeon, C.-Y. Wang, and M. Gastpar, "Computation over Gaussian networks with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7841–7861, Dec. 2014.
[12] C.-Y. Wang, S.-W. Jeon, and M. Gastpar, "Interactive computation of type-threshold functions in collocated Gaussian networks," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4765–4775, Sep. 2015.
[13] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+\text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.
[14] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.
[15] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.
[16] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6180–6195, Sep. 2011.
[17] M. Goldenbaum and S. Stanczak, "Robust analog function computation via wireless multiple-access channels," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3863–3877, Sep. 2013.
[18] M. Goldenbaum, H. Boche, and S. Stanczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4893–4906, Oct. 2013.
[19] A. Kortke, M. Godenbaum, and S. Stanczak, "Analog computation over the wireless channel: A proof of concept," in *Proc. IEEE Sens.*, Valencia, Spain, Nov. 2014, pp. 1224–1227.
[20] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 755–764, Apr. 2005.
[21] N. Ma, P. Ishwar, and P. Gupta, "Interactive source coding for function computation in collocated networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4289–4305, Jul. 2012.
[22] R. Appuswamy and M. Franceschetti, "Computing linear functions by linear coding over networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 422–431, Jan. 2014.
[23] S. Kannan and P. Viswanath, "Multi-terminal function multicasting in undirected graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 2334–2338.
[24] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York, NY, USA: Academic, 1981.
[25] S. M. Stigler, "Studies in the history of probability and statistics. XXXII: Laplace, Fisher and the discovery of the concept of sufficiency," *Biometrika*, vol. 60, pp. 439–445, Dec. 1973.

[26] O. Johnson, M. Aldridge, and R. Piechocki, "Delay-rate tradeoff in ergodic interference alignment," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 2626–2630.
[27] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3601–3621, Aug. 2006.
[28] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-aided relaying with adaptive link selection," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1530–1542, Aug. 2013.
[29] A. Jovičić, P. Viswanath, and S. R. Kulkarni, "Upper bounds to transport capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2555–2565, Nov. 2004.
[30] O. Lévêque and E. Telatar, "Information-theoretic upper bounds on the capacity of large extended ad hoc wireless networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 858–865, Mar. 2005.
[31] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3549–3572, Oct. 2007.
[32] B. Yu, J. Li, and Y. Li, "Distributed data aggregation scheduling in wireless sensor networks," in *Proc. IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2159–2167.
[33] J. Wieselthier, G. D. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010 pp. 585–594.

Sang-Woon Jeon (S'07–M'11) received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, South Korea, in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Deajeon, South Korea, in 2011. He has been an Assistant Professor with the Department of Information and Communication Engineering, Andong National University, Andong, South Korea, since 2013. From 2011 to 2013, he was a Postdoctoral Associate with the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale, Lausanne (EPFL), Switzerland. His research interests include network information theory and its application to wireless communications.

Dr. Jeon was the recipient of the Best Paper Award of the IEEE International Conference on Communications in 2015, the Best Thesis Award from the Department of EE, KAIST, in 2012, the Best Paper Award of the KICS Summer Conference in 2010, and the Bronze Prize of the Samsung Humantech Paper Awards in 2009.

Bang Chul Jung (S'02–M'08–SM'14) received the B.S. degree in electronics engineering from Ajou University, Suwon, South Korea, in 2002, and the M.S. and Ph.D. degrees in electrical and computer engineering from KAIST, Daejeon, Korea, in 2004 and 2008, respectively. He was a Senior Researcher/Research Professor with the KAIST Institute for Information Technology Convergence, Daejeon, South Korea, from January 2009 to February 2010. From March 2010 to August 2015, he was a Faculty member with Gyeongsang National University, Jinju, South Korea. He is currently an Associate Professor with the Department of Electronics Engineering, Chungnam National University, Daejon, South Korea. His research interests include 5G mobile communication systems, statistical signal processing, opportunistic communications, compressed sensing, interference management, interference alignment, random access, relaying techniques, device-to-device networks, in-network computation, and network coding.

Dr. Jung was the recipient of the Fifth IEEE Communication Society Asia-Pacific Outstanding Young Researcher Award in 2011. He was also the recipient of the Bronze Prize of Intel Student Paper Contest in 2005, the First Prize of KAIST's Invention Idea Contest in 2008, the Bronze Prize of Samsung Humantech Paper Contest in 2009, and the Outstanding Paper Award at the Spring Conference of Korea Institute of Information and Communication Engineering in 2015. He has been selected as a winner of the Haedong Young Scholar Award in 2015, which is sponsored by the Haedong Foundation and given by the Korea Institute of Communications and Information Science (KICS).