Energy Efficient Over-the-Air Computation for Correlated Data in Wireless Sensor Networks

Ayano Nakai-Kasai
Graduate School of Engineering
Nagoya Institute of Technology
Nagoya, Japan
nakai.ayano@nitech.ac.jp

Tadashi Wadayama
Graduate School of Engineering
Nagoya Institute of Technology
Nagoya, Japan
wadayama@nitech.ac.jp

Abstract—Over-the-air computation (AirComp) enables efficient wireless data aggregation in sensor networks by simultaneous processing of calculation and communication. This paper proposes a novel precoding method for AirComp that incorporates statistical properties of sensing data, spatial correlation and heterogeneous data correlation. The design of the proposed precoding matrix requires no iterative processes so that it can be realized with low computational costs. Moreover, this method provides dimensionality reduction of sensing data to reduce communication costs per sensor. We evaluate performance of the proposed method in terms of various system parameters. The results show the superiority of the proposed method to conventional non-iterative methods in cases where the number of receive antennas at the aggregator is less than that of the total transmit antennas at the sensors.

Index Terms—Over-the-air computation, wireless sensor networks, wireless data aggregation, dimensionality reduction

I. INTRODUCTION

In the fifth or more generation communication systems, one of the core technologies is to connect a large number of Internet-of-Things (IoT) devices that have abilities of sensing, computation, and wireless communication and to utilize their sensing data for many practical applications [1], [2]. There are a lot of active applications of sensor networks composed of the sensing devices, such as in agriculture [3] and in environmental monitoring [4], [5]. In a centralized data processing, data from distributed sensing devices are collected via wireless communication at an aggregator, which performs calculations to achieve desired actions for the applications. This procedure is called wireless data aggregation. The sensing data for such applications, for example, temperature, humidity, amount of chemicals, and soil conditions, often have spatial correlation and correlation among data types [6], [7]. It has been shown in the machine learning research area that the explicit use of the spatial correlation among sensors and the heterogeneous data correlation improves the processing accuracy [8], [9].

Efficient communication for sensor networks is also a hot research topic in the field of IoT wireless communications. Sensing devices usually have tiny batteries so that it is important for stable use of sensor networks to reduce power consumption related to wireless communication. In other words, the reduction of communication costs per sensor node is a crucial concern for the sensor networks. There are several studies that focus on dimensionality reduction of sensing data at the sensors [10]–[12] to reduce the frequency of communication per sensor node. It is desirable to design the reduction so as not to affect the desired operation as much as possible.

In addition, efficient data processing at the aggregator is required in IoT applications to achieve immediate response on demands in the next generation systems. The idea of over-the-air computation (AirComp) was first investigated in the field of information theory [13] and has been gathering much attention from a signal processing perspective in sensor networks [2], [14], which enables simultaneous processing of calculation of the sum of transmitted signals and communication at the aggregator by using analog-wave superposition property of wireless multiple-access channels (MAC). This is different from classical data aggregation settings where all the data are separately received and then the sum is calculated at the aggregator. The classical settings require significant bandwidth and cause heavy traffic, especially in large-scale sensor networks.

AirComp in sensor networks occurs aggregation errors, i.e., errors between the actual sum of transmitted data from the sensors and the aggregated data on the air, due to different channel coefficients of the sensors and additive noise at the aggregator. To reduce the aggregation errors, scaling coefficients or precoding matrices are applied to the transmit signals for the sensors. The optimization of them has been tackled in various contexts [15]–[18]. Huh et al. have proposed a gradient descent based optimization of the precoding matrix [15], but it requires high computational costs due to iterative calculation. The works [16]–[18] derived non-iterative optimization methods but [16] can be applied only to a scalar measurement, and [17] and [18] require dimensionality expansion of the sensing data. This expansion is not suitable because it requires high communication costs at the sensors. In addition, these methods do not consider correlation among the sensors which is potentially useful for improving the system performance.

In this paper, we propose a novel precoding scheme for AirComp in wireless correlated data aggregation. We introduce the spatial correlation and heterogeneous data correlation into the design of the precoding matrix. The proposed method...
is non-iterative and allows dimensionality reduction. These properties lead to energy efficient operation with low computational and communication costs in sensor networks. The contributions of this paper are summarized as follows:

- The proposed precoding matrix is derived by explicitly employing the correlation among sensors and data types. That correlation is not included in the conventional precoding methods for AirComp.
- The proposed method requires no iterative procedures. This is motivated by the idea in [19] and [20], which is proposed in a different context from AirComp. The computational cost is greatly less than the gradient descent based method [15].
- Dimensionality reduction is introduced to the proposed method. This helps lower communication costs and reduce battery consumption. The conventional non-iterative methods [17], [18] are not applicable to the dimensionality reduction.
- Simulation results on synthetic data show the superiority of the proposed method over other non-iterative solutions in cases where the number of receive antennas at the aggregator is less than that of the total transmit antennas at the sensors. This indicates that the proposed method is especially suitable for large-scale sensor networks.

II. PRELIMINARIES

A. Notations

In the rest of the paper, we use the following notation. Superscripts $\cdot^T$ and $\cdot^H$ denote the transpose and the Hermitian transpose, respectively. The zero vector, zero matrix and identity matrix are represented as $\mathbf{0}$, $\mathbf{O}$, and $\mathbf{I}$, respectively. The $\ell_2$-norm is $\| \cdot \|$. The complex circular Gaussian distribution $CN(\mathbf{m}, \mathbf{\Sigma})$ has mean vector $\mathbf{m}$ and covariance matrix $\mathbf{\Sigma}$. The expectation and trace operators are $E[\cdot]$ and $\text{Tr} [\cdot]$, respectively. We denote the set of complex block diagonal matrices with $k$ diagonal blocks of $m \times n$ matrices as $\mathbb{B}^{m \times n}_k$. The function $\alpha^+$ for $\alpha \in \mathbb{R}$ denotes $\max(0, \alpha)$. Hadamard product is represented as $\odot$, which is the element-wise multiplication of matrices.

B. System Model

Assume a wireless data aggregation system with a single aggregator with $r$ receive antennas and $K$ sensor nodes with $m$ transmit antennas per node as illustrated in Fig. 1.

Let $\mathbf{d}_k \in \mathbb{R}^n$ ($k = 1, \ldots, K$) be a vector composed of $n$ measurements at $k$th node and $\mathbf{d}_{kl} \in \mathbb{R}$ ($l = 1, \ldots, n$) be $l$th element of the vector. The size should be set to $m < n$ to apply dimensionality reduction, but the following method is not limited to this setting.

In many applications of sensor networks, the objective of the aggregator is to obtain some function value of sensors’ raw measurements. For example, arithmetic mean, weighted sum, or Euclidean norm is used as the function. Such functions have been named nomographic functions [14], [18] because they can be represented by combination of pre- and post-processing functions of the measurements. The nomographic function $f_k(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is applied to each element of the vectors and given by

$$f_k(d_{1\ell}, \ldots, d_{K\ell}) = \psi_k \left( \sum_{k=1}^{K} \varphi_{k\ell}(d_{k\ell}) \right),$$

(1)

where $\psi_k(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is a post-processing function and $\varphi_{k\ell}(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ is a pre-processing function, respectively. For example, the element-wise weighted sum $f_k(d_{1\ell}, \ldots, d_{K\ell}) = \sum_{k=1}^{K} \omega_{k\ell} d_{k\ell}$ can be represented by the pre-processing function $\varphi_{k\ell}(\chi) = \omega_{k\ell} \chi$ and the post-processing one $\psi_k(\chi) = \chi$.

From the form of the nomographic function (1), the aggregator wants to know the sum $\sum_{k=1}^{K} \varphi_{k\ell}(d_{k\ell})$ via communication with the nodes. We redefine the pre-processed local function values at node $k$ as a vector $\mathbf{x}_k = [\varphi_{k1}(d_{k1}), \ldots, \varphi_{kn}(d_{kn})]^T \in \mathbb{R}^n$ and then the element-wise sum can be summarized as

$$s = \sum_{k=1}^{K} \mathbf{x}_k \in \mathbb{R}^n.$$  

(2)

Assume that the further summarized vector $\mathbf{x} = [x_1^T, \ldots, x_K^T]^T \in \mathbb{R}^{nK}$ follows $\mathbf{x} \sim CN(\mathbf{0}, \mathbf{K})$. The covariance matrix $\mathbf{K} = E[\mathbf{xx}^T] \in \mathbb{R}^{nK \times nK}$ is positive definite and known to the aggregator. It includes information on the spatial correlation as the non-block diagonal elements and on the heterogeneous data correlation as the non-diagonal elements of each block.

Each node multiplies a precoding matrix $\mathbf{A}_k \in \mathbb{C}^{m \times n}$ by its own pre-processed vector $\mathbf{x}_k$ for reducing aggregation error, that is, the node $k$ transmits

$$c_k = \mathbf{A}_k \mathbf{x}_k \in \mathbb{C}^m$$

(3)

to the aggregator. The size $m$ is assumed to be smaller than that of $\mathbf{x}_k$ for dimensionality reduction. The aggregated signal through MAC is given by

$$y = H_1 c_1 + \cdots + H_K c_K + \mathbf{n} = \left( \sum_{k=1}^{K} H_k \mathbf{A}_k \mathbf{x}_k \right) + \mathbf{n} \in \mathbb{C}^r,$$  

(4)

where $H_k \in \mathbb{C}^{r \times m}$ is a channel matrix between node $k$ and the aggregator, and $\mathbf{n} \in \mathbb{C}^r$ is the additive noise that follows $\mathbf{n} \sim CN(\mathbf{0}, \mathbf{S})$. The positive definite covariance matrix $\mathbf{S} \in \mathbb{C}^{r \times r}$ is not limited to this setting.
\( C^{r \times r} \) represents correlation of the noise vector and is known to the aggregator. The model (4) is summarized as
\[
y = HAx + n, \tag{5}
\]
where
\[
H = [H_1, \ldots, H_K] \in C^{r \times mK}
\]
and
\[
A = \begin{bmatrix}
A_1 & 0 \\
0 & A_K
\end{bmatrix} \in C^{mK \times nK}.
\]

The aggregator is assumed to know information of the statistical properties of the transmit signal \( x \) and the noise \( n \), the received signal \( y \), and the channel matrices \( \{H_k\}_{k=1}^K \).

### III. PROPOSED METHOD

In this section, we describe how to design the proposed precoding matrix \( A \) that includes the correlation of data with non-iterative procedure.

#### A. Optimization Problem for Design of Proposed Precoder

The objective at the aggregator is to obtain the sum
\[
s = \sum_{k=1}^K x_k = Qx, \tag{6}
\]
where \( Q = [I, \ldots, I] \in \mathbb{R}^{n \times nK} \), as correctly as possible by using the available information because the received signal is distorted by the channel and noise. For this objective, we assume that the aggregator employs a linear MMSE estimate [21]
\[
\hat{s} = Wy = QKA^H H^H (HAKA^H H^H + S)^{-1} y. \tag{7}
\]
The matrix
\[
W = QKA^H H^H (HAKA^H H^H + S)^{-1} \in C^{n \times r}
\]
determined by minimizing the MSE:
\[
E[\| \hat{s} - s \|^2] = E[\| WHAx + Wn - Qx \|^2]. \tag{8}
\]

In this paper, we explore a matrix \( A \) that minimizes \( E[\| \hat{s} - s \|^2] \) of the linear MMSE estimation. Moreover, the limited power of the sensor nodes should be also taken into consideration for the design of the matrix \( A \). The total transmit power of the nodes is \( \sum_{k=1}^K E[\| A_k x_k \|^2] = E[\| Ax \|^2] \). Therefore, we consider the optimization problem:

\[
(P1) \quad \begin{aligned}
\min_A & E[\| \hat{s} - s \|^2] \\
\text{s.t.} & E[\| Ax \|^2] = P_0, \quad A \in \mathbb{B}^{m \times n}_K,
\end{aligned}
\]
where the first constraint means that the total transmit power is set to be \( P_0 > 0 \) and the second constraint means that \( A \) has a block diagonal structure. The problem \( P1 \) can be rewritten by substituting (7) into the cost function and results in an optimization problem of matrix function:

\[
(P2) \quad \begin{aligned}
\min_A & \text{Tr} \left[ Q \left( K^{-1} + A^H H^H S^{-1} H A \right)^{-1} Q^T \right] \\
\text{s.t.} & \text{Tr}[AKA^H] = P_0, \quad A \in \mathbb{B}^{m \times n}_K.
\end{aligned}
\]
The problem is nonconvex and difficult to solve in general. To make matters worse, the matrix \( A \) to be optimized has a block diagonal structure, which complicates the optimization process.

#### B. Policy

This paper proposes a closed-form solution of the non-convex problem \( P2 \) along with the idea in the conventional methods [19], [20] proposed in a different context from AirComp. We first consider the problem:

\[
\begin{aligned}
(P2') & \min_A \text{Tr} \left[ Q \left( K^{-1} + A^H H^H S^{-1} H A \right)^{-1} Q^T \right] \\
& \text{s.t.} \text{Tr}[AKA^H] = P_0.
\end{aligned}
\]

This is a relaxed problem of \( P2 \) where the block diagonal constraint is omitted. The problem \( P2' \) is known to have an optimal solution when the number of nodes is \( K = 1 \) [19], [20]. We obtain the solution \( \hat{A} \) of the problem \( P2' \) by using diagonalization of some matrices as employed in [19] and [20]. The solution is not exactly optimal when \( K > 1 \) but it gives us a close formula of the preceding matrix. Next, a block diagonal matrix \( \hat{A}_{bd} \) is derived by removing non-block diagonal elements from \( \hat{A} \). This strategy has been employed in [22] for not-AirComp settings. Finally, we obtain the solution \( \hat{A} \) by scaling the norm of \( \hat{A}_{bd} \) to satisfy the power constraint \( \text{Tr}[AKA^H] = P_0 \).

#### C. Derivation of Non-Blockdiagonal matrix \( \hat{A} \)

This section derives non-block diagonal solution \( \hat{A} \) by diagonalization of matrices in the objective function. We describe eigenvalue decomposition of the matrices \( K \) and \( H^H S^{-1} H \) as
\[
K = U \Delta U^H \in \mathbb{R}^{nK \times nK}, \tag{13}
\]
\[
H^H S^{-1} H = V \Lambda V^H \in \mathbb{C}^{mK \times mK}, \tag{14}
\]
respectively, where \( U \in \mathbb{R}^{nK \times nK}, \ V \in \mathbb{C}^{mK \times mK} \) are unitary matrices, and \( \Delta \in \mathbb{R}^{nK \times nK}, \ \Lambda \in \mathbb{R}^{mK \times mK} \) are diagonal matrices. The diagonal elements are represented as
\[
\Delta = \begin{bmatrix}
\delta_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \delta_{nK}
\end{bmatrix}, \quad \Lambda = \begin{bmatrix}
\lambda_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \lambda_{mK}
\end{bmatrix},
\]
respectively. Without loss of generality, we assume \( \delta_1 \geq \delta_2 \geq \cdots \geq \delta_{nK} \geq 0 \) and \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{mK} \geq 0 \). When \( r < mK \), the matrix \( \Lambda \) has the property \( \lambda_{r+1} = \cdots = \lambda_{mK} = 0 \). We further assume that the matrix variable \( A \) is decomposed as
\[
A = V \Phi U^H, \quad \Phi = \begin{bmatrix}
\phi_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \phi_{mK}
\end{bmatrix}, \tag{15}
\]
where \( \{\phi_j\}_{j=1}^{mK} \) are scalar parameters and satisfy \( |\phi_j|^2 \geq 0 \). That decomposed formulation is motivated by the work [19], [20]. Therefore, the solution of the problem \( P2' \) is assumed to be fully determined by the \( mK \) parameters.
From these diagonalized representations, we can solve the problem $P2'$ in terms of the parameters $\{\phi_j\}_{j=1}^{mK}$.

**Theorem 1:** The solution of the problem $P2'$ is given as the following water-filling problem

\[
|\phi_j|^2 = \begin{cases} 
\frac{1}{\delta_j \lambda_j} \left( \sqrt{\delta_j \lambda_j R_j / \mu} - 1 \right)^+, & (j = 1, \ldots, \min(r, mK)), \\
0, & (j = \min(r, mK) + 1, \ldots, mK),
\end{cases}
\]

where $R_j = \sum_{i=1}^{n} (QU)_{ij}^2$, $QU$ is the $(i, j)$th element of $QU$, and $\mu \in \mathbb{R}$ is determined to satisfy the power constraint in the problem. For the case of $j = 1, \ldots, \min(r, mK)$, if the elements in $\cdot^+$ in the right-hand side of (16) are nonnegative for all $j$, the solution is given by

\[
\hat{\phi}_j = \left( \frac{1}{\delta_j \lambda_j} \left( \sqrt{\delta_j \lambda_j R_j / \mu} - 1 \right)^+ \right) \\
\frac{1}{\sum_{j=1}^{\min(r, mK)} \frac{1}{\lambda_j}}.
\]

**Proof:** We represent the cost function of the problem $P2'$ as $f(\{\phi_j\}_{j=1}^{mK})$. The function can be rewritten as

\[
f(\{\phi_j\}_{j=1}^{mK}) = \text{Tr} \left[ QU \left( \Delta^{-1} + \Phi^H \Lambda \Phi \right)^{-1} U^H Q^T \right]
\]

by using (13)-(15). Note that the matrices $Q$ and $U$ are known at the aggregator and then it is possible to calculate the trace in (18) directly. The result of expanding the equation is

\[
f(\{\phi_j\}_{j=1}^{mK}) = \sum_{j=1}^{\min(r, mK)} \frac{\delta_j R_j}{1 + \delta_j \lambda_j |\phi_j|^2} + \sum_{j=\min(r, mK)+1}^{mK} \delta_j R_j.
\]

On the other hand, the transmit power $\text{Tr}[AKA^H]$ in the problem $P2'$ is also given by

\[
\text{Tr}[AKA^H] = \text{Tr}[\Phi \Delta \Phi^H] = \sum_{j=1}^{mK} \delta_j |\phi_j|^2.
\]

In order to obtain the solution of the problem $P2'$, we set the following Lagrangian function

\[
\mathcal{L}(\{\phi_j\}_{j=1}^{mK}, \mu) = f(\{\phi_j\}_{j=1}^{mK}) - \mu \left( P_0 - \sum_{j=1}^{mK} \delta_j |\phi_j|^2 \right),
\]

where $\mu$ is the Lagrange multiplier. The condition in this case is \(\frac{\partial \mathcal{L}}{\partial \phi_j} = \frac{\partial \mathcal{L}}{\partial \mu} = 0\). By solving \(\frac{\partial \mathcal{L}}{\partial |\phi_j|^2} = 0\) in terms of $|\phi_j|^2$, we can obtain

\[
|\phi_j|^2 = \begin{cases} 
\frac{1}{\delta_j \lambda_j} \left( \sqrt{\delta_j \lambda_j R_j / \mu} - 1 \right)^+, & (j = 1, \ldots, \min(r, mK)), \\
0, & (j = \min(r, mK) + 1, \ldots, mK),
\end{cases}
\]

(22)

where the function $\cdot^+$ is applied because $|\phi_j|^2 \geq 0$.

If the arguments in the right-hand side of (22) become nonnegative for all indeces $j$, another relation can be derived by substituting (22) into $\frac{\partial \mathcal{L}}{\partial \mu} = 0$ and we then have

\[
\frac{1}{\sqrt{\mu}} = \frac{P_0 + \sum_{j=1}^{\min(r, mK)} \frac{1}{\lambda_j}}{\sum_{j=1}^{\min(r, mK)} \sqrt{\delta_j R_j / \lambda_j}}.
\]

(23)

From (22) and (23), we can obtain the solution $\hat{\phi}_j$.

If there exists index $j$ where the argument in the right-hand side of (22) becomes negative, the Lagrange multiplier $\mu$ is determined by the well-known water-filling algorithm (Sect. 3.E in [23]) to satisfy the power constraint.

We can obtain the matrix $\tilde{A}$ by using the solution $\{\phi_j\}_{j=1}^{mK}$ and constructing from (15).

\[\tilde{A}_{bd} = M \otimes \tilde{A}.\]

(24)

D. Block Diagonalization

The problem we should solve is $P2$ and the matrix $A$ must have block-diagonal structure. In this paper, we omit the non-block diagonal elements of the matrix $\tilde{A}$ obtained in the previous section and then rescale to the constrained power.

Let $M \in \mathbb{R}^{m \times n}$ be a masking matrix composed of $K \times K$ blocks where the diagonal blocks are the matrices whose components are all 1 and the non-diagonal blocks are all 0. The block diagonalized matrix $\tilde{A}_{bd}$ can be represented as the element-wise multiplication of $M$ and $\tilde{A}$, i.e.,

\[\tilde{A}_{bd} = M \otimes \tilde{A}.\]

(25)

Note that the matrix $\tilde{A}$ satisfies the power constraint $\text{Tr}[\tilde{A}_{bd}^H \tilde{A}_{bd}] = P_0$ because of the constraint of the problem $P2'$ but the block diagonalized matrix $\tilde{A}_{bd}$ does not. Therefore, we rescale the norm of the matrix $\tilde{A}_{bd}$ to satisfy the power constraint. We then obtain the final solution

\[\tilde{A}_{bd} = \sqrt{\frac{P_0}{\text{Tr}[\tilde{A}_{bd}^H \tilde{A}_{bd}]}} \tilde{A}_{bd}.\]

(25)

The proposed method does not include iterative processes and the main factor of the computational costs is eigenvalue decomposition (13), which requires $O((nK)^3)$. However, this can be pre-processed as long as the covariance matrix $K$ is fixed, so that the complexity is much less than the gradient descent based method [15].

IV. SIMULATION RESULTS

Performance of the proposed method was evaluated via computer simulations. We evaluated influence of system parameters on the averaged and normalized squared error, i.e.,

\[
\frac{\sum_{i=1}^{T} \sum_{z=1}^{Z} \| \hat{s}_{iz} - s_{iz} \|^2 / (nKTZ)}{\sum_{i=1}^{T} \sum_{z=1}^{Z} \| s_{iz} \|^2}
\]

where $T$ is the number of generations of $H$, $Z$ is the number of generations of $x$ for a single generation of $H$, and $\hat{s}_{iz}$ and $s_{iz}$ are corresponding instances. We set to $T = 10$ or more and $Z = 100$. Specifically, the simulation results are examined in terms of the performance dependency on

- Data compression ratio $mK/nK$: ratio of the total number of transmit antennas and that of measurements,
- Communication compression ratio $r/mK$: ratio of the number of receive antennas and the total number of transmit antennas,
- The number of nodes $K$,
- and Signal-to-noise ratio (SNR) (dB).

In this paper, we define the SNR as $10 \log_{10}(P_0/\text{Tr}[S])$ (dB). The length of the measurement was set to $n = 8$ and the total transmit power was $P_0 = 10$ in all the simulations.
covariance matrices $K$ and $S$ had correlated formulations and the elements were determined as

$$[K]_{ij} = 0.8^{|i-j|}, \quad [S]_{ab} = 0.5^{|a-b|}/r,$$

for $i,j = 1,\ldots,n_K$ and $a,b = 1,\ldots,r$, respectively. Each element of the channel matrix $H$ was identically and independently generated by $CN(0,1)$. Moreover, we compared the performance of the proposed method with the following three schemes:

1) Communicate-then-Compute: method where optimization is done in the same manner as the proposed method but the objective function is not for $s$ but $x$ (normal MMSE estimation, not specifically designed for AirComp),

2) Random: method using a random matrix as $A$ where each element is identically and independently generated by $CN(0,1)$ and normalized to satisfy the power constraint. Such a dimensional reduction is typically employed in some estimation methods [24],

3) and Huh et al.: iterative gradient descent based method [15] for deriving the matrices $A_k$ from the problem $P1$. The method requires higher computational cost than the proposed and other methods so that it is regarded as a baseline.

The number of iterations for the method [15] was set to 10.

Fig. 2 shows the evaluation with respect to the data compression ratio. The system parameters were set to $(n, K, r, SNR(dB)) = (8, 30, 5m, 25)$. From the figure, the iterative method of Huh et al. achieves the lowest error at the cost of high computational costs. The proposed method shows the best performance at any data compression ratio among the non-iterative methods.

The key feature of the proposed method is revealed from viewpoints of communication compression ratio and the number of nodes, shown in Fig. 3 and Fig. 4, respectively.

The evaluation in terms of communication compression ratio is shown in Fig. 3 where the system parameters were set to $(m, K, SNR(dB)) = (2, 30, 25)$. The MSE curves of the proposed and communicate-then-compute methods appear to have two modes. In the region $1 \leq r/mK$, i.e., when the number of the receive antennas $r$ is sufficiently large, the communicate-then-compute method shows lower error than the proposed method. On the other hand, in the region $r/mK < 1$, the proposed method achieves lower error than the other methods.

In Fig. 4 we evaluated the influence of the number $K$ of nodes on normalized MSE. The system parameters were set to $(m, r, SNR(dB)) = (2, 16, 25)$. In all the methods, the smaller the number of nodes, the lower the estimation error. The error of the proposed method is the lowest and the performance difference from the other methods becomes larger with increase of $K$. These results indicate that the proposed method is suitable in situations where the number of receive antennas are limited but the number of nodes is increasing. The situations are nothing short of typical IoT environments.

Finally, we evaluated performance of the proposed method in different SNR. The system parameters were set to $(n, m, r) = (8, 2, 16)$ and $(n, K, r, SNR(dB)) = (8, 30, 5m, 25)$. From the figure, the proposed method shows the best performance at any data compression ratio among the non-iterative methods.

Fig. 2. MSE vs. data compression ratio where $(n, K) = (8, 30)$, $r = 5m$, and SNR= 25(dB).

Fig. 3. MSE vs. communication compression ratio where $(n, m, K) = (8, 2, 30)$ and SNR= 25(dB).

Fig. 4. MSE vs. the number of nodes where $(n, m, r) = (8, 2, 16)$ and SNR= 25(dB).
proposed method involving more sophisticated operations of IoT environments. Future work includes the extension of the proposed method to achieve better performance in typical sensor networks. In other words, the number of transmit antennas at the sensors is less than the total number of receive antennas at the aggregator is less than the total number of transmit antennas at the sensors. In other words, the proposed method achieves better performance in typical sensor networks. Moreover, this method is suitable for sensor networks with limited batteries because it includes dimensionality reduction of the transmit vectors and can reduce communication costs per sensor. These properties enable energy efficient wireless data aggregation. Simulation results on synthetic data showed that the performance of the proposed method is the best among the methods at any SNR.

V. CONCLUSIONS

This paper has proposed a novel precoding method for Aircomp in wireless data aggregation. The proposed precoding matrix has been derived by explicitly employing the correlation of data, which is appeared in typical applications of sensor networks. Moreover, the proposed method has a non-iterative form so that it does not require high computational costs. Furthermore, this method is suitable for sensor networks with limited batteries because it includes dimensionality reduction of the transmit vectors and can reduce communication costs per sensor. These properties enable energy efficient wireless data aggregation. Simulation results on synthetic data showed that the performance of the proposed method is better than other methods, especially when the number of receive antennas at the aggregator is less than the total number of transmit antennas at the sensors. In other words, the proposed method achieves better performance in typical IoT environments. Future work includes the extension of the proposed method involving more sophisticated operations of block diagonalization.

REFERENCES

[1] T. Salam, W. U. Rahman, and X. Tao, “Data aggregation in massive machine type communication: Challenges and solutions,” IEEE Access, vol. 7, pp. 41921–41946, Apr. 2019.
[2] G. Zhu, J. Xu, K. Huang, and S. Cui, “Over-the-air computing for wireless data aggregation in massive IoT,” IEEE Wireless Commun., vol. 28, no. 4, pp. 57–65, Aug. 2021.
[3] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and MHD N. Hindia, “An overview of Internet of things (IoT) and data analytics in agriculture: Benefits and challenges,” IEEE Internet Things J., vol. 5, no. 5, pp. 3758–3773, Oct. 2018.
[4] V. Šmilď and R. Hofman, “Tracking of atmospheric release of pollution using unmanned aerial vehicles,” Atmospheric Environment, vol. 67, pp. 425–436, Mar. 2013.
[5] M. F. Othman and K. Shazali, “Wireless sensor network applications: A study in environment monitoring system,” Procedia Engineering, vol. 41, pp. 1204–1210, 2012.

\[(m, r, K) = (2, 30, 30).\] From Fig. 5, the performance of the proposed method is the best among the methods at any SNR.