The following publication Q. Wang, S. Guo and K. -F. C. Yiu, "Distributed Acoustic Beamforming With Blockchain Protection," in IEEE Transactions on Industrial Informatics, vol. 16, no. 11, pp. 7126-7135 is available at https://doi.org/10.1109/TII.2020.2975899.

Distributed Acoustic Beamforming with Blockchain Protection

Qingzheng Wang, Shan Guo, Ka-fai Cedric Yiu,

***Abstract*—Speech is a natural user interface for the Internet of Things system. However, the presence of noise affects severely the performance of such system. With the deployment of smart devices with microphones, one can form a powerful acoustic sensor network to enhance the speech via beamforming techniques. On the other hand, reliability of data transmission also determines the beamforming performance, since faulty data will drift the beamformer steering location randomly. Currently there is no protection scheme for acoustic data transmitted over the wireless network in order to keep steady beamforming performance. In this paper, we design a compound distributed beamformer where nodes are grouped and the system are embedded with blockchain technology to protect the data integrity during transmission. It attempts to provide more possible reliable connections between groups. Simulated experiments shows that the distributed beamformer with blockchain protection is able to maintain steady beamforming performance.**

***Index Terms*—Signal enhancement; Distributed acoustic array network; Blockchain data protection.**

I. Introduction

THE Internet of Things (IoT) has transformed our daily life in many aspects. By providing a seamless integration of physical objects into the information network [1], this enables information exchange to flow between devices and also allows the devices to be controlled remotely by users via a range of man-to-machine speech interactive systems [2]. People can remotely control the IoT devices using voice commands or natural dialogue [3]. Voice control is an attractive feature that provides a natural mean of remote control and becomes the primary user interface for the smart home [4]. However, the interference of the environment and background noises degrade the performance of such devices [5]. In order to have smooth operations, acoustic noise should be suppressed and the required speech to be enhanced.

With the advent of wireless smart devices equipped with microphones, a wireless acoustic sensor network (WASN) can be formed and many innovative applications can be developed. One important application is to enhance speech signals and suppress unwanted noise via beamforming techniques [6], [7]. If successful, this can enhance significantly the capability of voice control device. Since the microphone array in WASN no longer needs to be wired in a restricted area as the traditional microphone array does but can be placed in any suitable position, a WASN could accommodate many sensor nodes which are positioned anywhere and each node is allowed to contain a microphone array rather than a single microphone. There are several challenges in developing the distributed beamforming system. First, although the sensor coverage becomes larger and speech signal can be enhanced with the increment of microphone arrays, the increased computational burden cannot be ignored. Second, reliability of data transmission determines the performance of the designed beamformers, since faulty data will drift the filter coefficients quickly and deviated from the target location. In order to enhance robustness and reliability of voice control system, the transmission reliability via wireless channel should be dealt with. In this paper, we provide an innovative solution including two major parts:

- A novel beamforming technique which distributes the computational burden over the nodes of the WASN;
- A novel data protection scheme in which blockchain technique ensures the integrity of transmitted data between the nodes of the WASN.

When it comes to the designing of beamformer systems, the key step is to compute the vector of beamforming weight which is the solution of an optimization problem. In this procedure, there are two important aspects. One is to decide the evaluation criterion of the optimization problem, i.e., the objective function. There are commonly two choices: minimum variance distortionless response (MVDR) [8] and minimum mean square error (MMSE) [9]. While the MVDR beamformer requires an exact steering vector, the MMSE beamformer makes use of reference signals in the design. The other important aspect is to decide the transmission and processing strategy. There are also two typical choices: the centralized beamformer and distributed beamformer. In the centralized beamformer, all raw acoustic signals are collected in a fusion center which conceptually connects to all the acoustic sensors [10]. The fusion center calculates the inverse covariance matrix of acoustic signals so as to derive an optimal beamforming weight. However, the centralized beamformer may not be suitable for the WASN because the total number of microphones is too large to process in a signal device when it is employed in the WASN [11]. Apart from that, the centralized beamformer is limited by the communication bandwidth and transmission power [12]. Moreover, the fusion center could be absent in the WASN due to the uncertainty topology of the wireless network [13], [14]. Last but not least, the transmission failure between the fusion center and acoustic sensors cannot be ignored in the WASN. To avoid these shortcomings, the distributed beamformer [15] is widely adopted recently. Since each node in the WASN has its own processing unit, it can locally process data and share the results with their neighboring nodes. By cooperating in a distributed

Q. Wang, S. Guo and K.F.C. Yiu are with the Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China; e-mail: cedric.yiu@polyu.edu.hk

fashion, all the nodes share the computational burden.

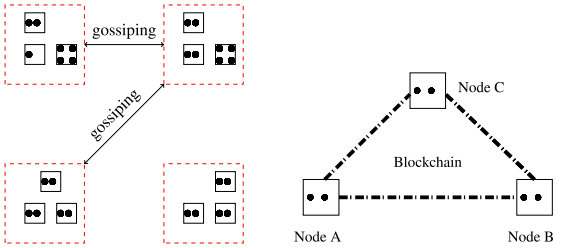
Although a distributed beamforming system has advantages, it has also created new challenges. One challenge is how to design decentralized schemes so that the computational burden is shared by all the nodes and the entire sampling data is fused iteratively in the WASN. A good decentralized scheme should reduce the wireless data transmission and share the computational burden between nodes. A newly proposed strategy of data transmission for distributed beamformers is called the gossip algorithm where information exchanges between adjacent nodes constantly and successively [13]. With successive iterations, this algorithm can reach the consensus solution for each node [16]. There are variants of the gossip algorithm and will be discussed in detail in Section III. The gossip algorithm has several advantages. First of all, since only one node will communicate with one of its neighbors at each iteration, it is computationally efficient. Second, the algorithm does not require the WASN to remain to be the same throughout the whole process, but allowing new links to append and old ones to exit. However, there exist challenges in the implementation of gossip algorithms. When the number of nodes increases, the number of iterations required to reach convergence will increase rapidly [17]. To cope with this drawback, our approach only employs the gossip method among groups to trade off the increasing complexity. In addition, convergence of the algorithms can be slowed down significantly due to faulty transmission caused by the unstable links in the WASN [18], which in turns degrades the overall performance of the system. Moreover, it is necessary to deal with the consensus issue of nodes inside same groups caused by faulty transmission. Thus, a data protection mechanism is very important for the sake of data integrity.

For multimedia applications, tranmission is often carried out via protocols like user datagram protocol (UDP) [19] to increase efficiency; on the other hand, sacrificing data integrity with less verification [20]. Unstable wireless links may yields heavy packet loss. In order to retain transmission reliability, it has must be carried out within the application level. First of all, corrupted data should be rejected or discarded by receivers. There are various ways of detecting faulty transmission. The simplest method is Cyclic Redundancy Check (CRC) but it is limited by its error detection capability [21]. A more elaborated method is called Message Authentication Code (MAC) based on a cryptographic-based algorithm [22], [23]. A lightweight data integrity checking method is to use the watermarking technique upon sensor networks [24], [25]. However, most of existing integrity mechanisms for wireless networks require a base station (or fusion center) which is likely absent for most distributed wireless networks. In addition, this centralized architecture has certain inherent vulnerabilities. For example, the whole system stops working if the base station is down due to maintenance or software failures [26]. Furthermore, the aforementioned methods only detect the corruption of transmitted data rather than improving the data integrity. Transmission Control Protocol (TCP) has been extended and adapted to be deployed in wireless sensor network so that data can be retransmitted to improve the data integrity [27]. However, it relies on the original established links. If one wireless link has become unreliable, correct data still cannot be obtained by the retransmission request.

In view of the above, here we design a data protection scheme in application level for the WASN using blockchain technique. We propose a novel framework of the compound distributed beamformer where nodes are grouped and the system are embedded with blockchain technology [28] to protect the data during transmission. The distributed MMSE beamforming algorithm is developed. The sketch of the basic idea is shown in Fig. 1, where small black rectangles represent nodes in the WASN, and each node could contain several microphones which are represented by black dots. Nodes in the WASN are divided into groups, represented by the dotted red line according to certain preset rules. This is a two-level communication scheme containing the intra-group data sharing based on blockchain technique as well as the inter-group data communication via gossip algorithms. We first share data within each group and use the blockchain technique to protect the fused data as well as to resolve the consensus problem inside the group. Using the hash function, another group can easily verify the correctness of the received data in inter-group communications. In our proposed framework, we actually can randomly select any node in the group to establish the wireless link for data transmission between groups, since any selected node will have the same data within the group. If the selected link has a problem, we can immediately switch to another link so that we will not lose access to the entire group when one wireless link becomes unreliable. In our proposed framework, the connectivity reliability between groups is enhanced by providing more than one wireless link such that the possibility of faulty transmission is decreased.

Blockchain is a distributed storage system in which data is stored in a decentralized network as blocks and updated using an append-only structure. After the first introduction in 2008 by Satoshi Nakamoto, blockchain is growing with fast popularity [29]. It has been employed successfully in cryptocurrency and some other industries as well. Optimizations of blockchain have been conducted in the resource constrained environment [26]. In the design of the blockchain implementation for beamforming, we need to consider two important aspects including the reduction of computational complexity and the restriction of ledger scalability. In order to reduce the computational complexity, the distributed trust method is employed here to replace proof-of-work [30], since it decreases new block processing overhead while maintaining most of its security benefits. To deal with the problem of scalability, we first employ the short-time Fourier transform to locally compress raw acoustic signals in each node. It reduces the requirement of memory and bandwidth in the system. Second, we create new ledgers for each time frame and delete them after the computation of beamforming weights. Therefore, the length of ledger is bounded by the maximum iteration of gossip algorithms.

The rest of the paper is organized as follows. The problem formulation is given in Section II. The distributed computation scheme is introduced in Section III. The data protection based on blockchain technique is illustrated in Section IV. The simulation study is demonstrated in Section V. The discussion



(a) Gossiping among groups (b) Blockchain structure of the left bottom group in (a)

Fig. 1. Sketch of the proposed framework

on experimental results is presented in Section VI. Conclusion and further work are shown in Section VII.

II. Notation and Problem Formulation

In this work we consider an enclosed room with acoustic reverberation. In this room, N speech sources are settled at $\gamma_n, n = 0, \ldots, N-1$ and M-elements microphone array settled at $\delta_m, m = 1, \ldots, M$. The M microphones are grouped in U nodes. For the u-th node, it contains M_u microphones such that $M = \sum_{u=1}^{U} M_u$. In addition, we divide all the nodes into V groups. Each group contains U_v nodes with $U = \sum_{v=1}^{V} U_v$. Without loss of generality, the sensor at γ_0 is denoted as the signal of interest; the others are interferences; and the microphone at δ_1 is the reference microphone. Here, the noise placement information is not considered in beamformer design. Given the room dimension, sound speed, locations of sources and microphones, the time domain room impulse responses (RIR) $h(\delta_m, \gamma_n)$ from n-th source to m-th microphone can be generated by the image method [31]. Let $\mathbf{s}_n$ denote the signal at the source γ_n. The received signal at microphone δ_m is calculated by

$$\mathbf{s}_{m,n} = \mathbf{h}(\delta_m, \gamma_n) \odot \mathbf{s}_n \tag{1}$$

where $\odot$ denotes the convolution operator. By the short-time Fourier transform (STFT), the frequency domain coefficient of the m-th microphone is given by

$$Y_m(f,k) = S_{m,0}(f,k) + \sum_{n=1}^{N-1} S_{m,n}(f,k) + N_m(f,k) \tag{2}$$

where $S_{m,n}(f,k)$ is the STFT coefficient of $\mathbf{s}_{m,n}$ at frequency-bin index f and time-frame index k. The target source is $S_{m,0}(f,k)$. The interference sources are $\sum_{n=1}^{N-1} S_{m,n}(f,k)$. The noise STFT coefficient of m-th microphone is denoted by $N_m(f,k)$. Let $\mathbf{Y}(f,k) = [Y_1(f,k), \; Y_2(f,k), \; \ldots, \; Y_M(f,k)]^{\mathtt{T}}$ and $\mathbf{S}(f,k) = [S_{1,0}(f,k), \; S_{2,0}(f,k), \; \ldots, \; S_{M,0}(f,k)]^{\mathtt{T}}$. Here, $\mathbf{Y}(f,k)$ is the input data of the beamformer in operation phase, $\mathbf{S}(f,k)$ is the target signal. Let $w_m(f)$ be the beamforming weight of the m-th microphone at frequency f. The weight vector of the beamformer with frequency f is denoted as $\mathbf{w}(f) = [w_1(f) \; w_2(f) \; \cdots \; w_M(f)]^{\mathtt{T}}$. At the time-frame k and frequency f, the output of the beamformer in the frequency domain is defined by

$$\tilde{Y}(f,k) = \sum_{m=1}^{M} w_m(f) Y_m(f,k) = \mathbf{w}(f)^{\mathtt{H}} \mathbf{Y}(f,k). \tag{3}$$

Then, the problem of MMSE beamformer [9] can be recognized as the least square optimization problem which is formulated as

$$\mathbf{w}_{\mathtt{opt}}(f) = \arg\min_{\mathbf{w}(f)} \; \mathbb{E}\left\{ |\tilde{Y}(f,k) - S_r(f,k)|^2 \right\} \tag{4}$$

where $|\cdot|$ denotes the absolute value and $\mathbb{E}$ denotes the expectation operator, and $S_r(f,k)$ denotes the STFT coefficient related to the observation from the reference microphone. Let K_1 and K_2 denote the time-frame length in the optimization and operation phase respectively. Assume the reference signal $S_r(f,k)$ is independent of the actual observation $\mathbf{Y}(f,k)$, substituting Equation (3) into Equation (4), the original problem can be expressed as

$$\begin{aligned}
\mathbf{w}_{\mathtt{opt}}(f) &= \arg\min_{\mathbf{w}(f)} \left\{ \sum_{k=0}^{K_1-1} \left[|\mathbf{w}(f)^{\mathtt{H}} \mathbf{S}(f,k) - S_r(f,k)|^2 \right] \right. \\
&\qquad \left. + \sum_{k=0}^{K_2-1} |\mathbf{w}(f)^{\mathtt{H}} \mathbf{Y}(f,k)|^2 \right\} \\
&= \arg\min_{\mathbf{w}(f)} \left\{ \mathbf{w}(f)^{\mathtt{H}} [\hat{\mathbf{R}}_{SS}(f,K_1) + \hat{\mathbf{R}}_{YY}(f,K_2)] \mathbf{w}(f) \right. \\
&\qquad \left. - \mathbf{w}(f)^{\mathtt{H}} \hat{\mathbf{r}}_s(f,K_1) - \hat{\mathbf{r}}_s^{\mathtt{H}}(f,K_1) \mathbf{w}(f) + \hat{r}_{s_r} \right\}
\end{aligned}$$

where $\hat{r}_{s_r}$ is the variance of the reference microphone which can be treated as a constant in this optimization problem. In the optimization phase, the estimated correlation matrix $\hat{\mathbf{R}}_{SS}(f,K_1)$ and cross correlation vector $\hat{\mathbf{r}}_s(f,K_1)$ are calculated by

$$\hat{\mathbf{R}}_{SS}(f,K_1) = \frac{1}{K_1} \sum_{k=0}^{K_1-1} \mathbf{S}(f,k) \mathbf{S}(f,k)^{\mathtt{H}}, \tag{5}$$

$$\hat{\mathbf{r}}_s(f,K_1) = \frac{1}{K_1} \sum_{k=0}^{K_1-1} \mathbf{S}(f,k) S_r(f,k)^* \tag{6}$$

where $S_r(f,k)^*$ is the complex conjugate of $S_r(f,k)$. In the operation phase, the estimated correlation matrix $\hat{\mathbf{R}}_{YY}(f,K_2)$ is defined by

$$\hat{\mathbf{R}}_{YY}(f,K_2) = \frac{1}{K_2} \sum_{k=0}^{K_2-1} \lambda^{K_2-1-k} \mathbf{Y}(f,k) \mathbf{Y}(f,k)^{\mathtt{H}} \tag{7}$$

where λ is an exponential weighting factor.

Given the estimates from Equation (5)-(7), the optimal weight vector of the MMSE beamformer in Equation (4) is obtained by

$$\mathbf{w}_{\mathtt{opt}}(f) = \hat{\mathbf{R}}(f,K_2)^{-1} \hat{\mathbf{r}}_s(f,K_1) \tag{8}$$

where $\hat{\mathbf{R}}(f,K_2) = \hat{\mathbf{R}}_{SS}(f,K_1) + \hat{\mathbf{R}}_{YY}(f,K_2)$.

In time-frame k of the operation phase, the known information is all the observations in the optimization phase as well

as the observations up to time-frame k in the operation phase. Then, we have

$$\hat{\mathbf{R}}(f,k) = \hat{\mathbf{R}}_{SS}(f,K_1) + \hat{\mathbf{R}}_{YY}(f,k)$$

which can be extended to

$$\begin{aligned}\hat{\mathbf{R}}(f,k) &= \hat{\mathbf{R}}_{SS}(f,K_1) + \hat{\mathbf{R}}_{YY}(f,k)\\ &= \hat{\mathbf{R}}_{SS}(f,K_1) + \lambda\hat{\mathbf{R}}_{YY}(f,k-1) + \mathbf{Y}(f,k)\mathbf{Y}(f,k)^{\mathtt{H}}\\ &= \lambda\hat{\mathbf{R}}(f,k-1) + \mathbf{Y}(f,k)\mathbf{Y}(f,k)^{\mathtt{H}} + (1-\lambda)\hat{\mathbf{R}}_{SS}(f,K_1)\\ &= \lambda\hat{\mathbf{R}}(f,k-1) + \mathbf{Y}(f,k)\mathbf{Y}(f,k)^{\mathtt{H}}\\ &\quad + \sum_{m=1}^{M}(1-\lambda)\gamma_m(f)\mathbf{q}_m(f)\mathbf{q}_m(f)^{\mathtt{H}}\end{aligned}$$

where $\gamma_m(f)$ is the m-th eigenvalue and $\mathbf{q}_m(f)$ is the m-th eigenvector of the $M \times M$ correlation matrix $\hat{\mathbf{R}}_{SS}(f,K_1)$. With the use of a rank-one approximation of the matrix [32], $\hat{\mathbf{R}}(f,k)$ can be updated by

$$\begin{aligned}\hat{\mathbf{R}}(f,k) =& \lambda\hat{\mathbf{R}}(f,k-1) + \mathbf{Y}(f,k)\mathbf{Y}(f,k)^{\mathtt{H}}\\ &+ (1-\lambda)\gamma_i(f)\mathbf{q}_i(f)\mathbf{q}_i(f)^{\mathtt{H}}\end{aligned} \tag{9}$$

where $i = (k \bmod M) + 1$.

Using the Matrix Inversion Lemma [33] twice, the inverse correlation matrix $\hat{\mathbf{R}}(f,k)^{-1}$ can be computed iteratively

$$\hat{\mathbf{R}}(f,k)^{-1} = \check{\mathbf{R}}(f,k) - \frac{\gamma_i(f)(1-\lambda)\check{\mathbf{R}}(f,k)\mathbf{q}_i(f)\mathbf{q}_i(f)^{\mathtt{H}}\check{\mathbf{R}}(f,k)}{1+\gamma_i(f)(1-\lambda)\mathbf{q}_i(f)^{\mathtt{H}}\check{\mathbf{R}}(f,k)\mathbf{q}_i(f)} \tag{10}$$

where

$$\begin{aligned}\check{\mathbf{R}}(f,k) &= \lambda^{-1}\hat{\mathbf{R}}(f,k-1)^{-1}\\ &-\frac{\lambda^{-2}\hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k)\mathbf{Y}(f,k)^{\mathtt{H}}\hat{\mathbf{R}}(f,k-1)^{-1}}{1+\lambda^{-1}\mathbf{Y}(f,k)^{\mathtt{H}}\hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k)}.\end{aligned} \tag{11}$$

In order to reduce the influence of the random environmental noise, a first order autoregressive smoothing model is used to iteratively update the weight vector of the beamformer as

$$\mathbf{w}^{\mathtt{k}}(f) = \alpha\mathbf{w}^{\mathtt{k-1}}(f) + (1-\alpha)\hat{\mathbf{R}}(f,k)^{-1}\hat{\mathbf{r}}_s(f,K_1) \tag{12}$$

where $\alpha \in (0,1)$ is the smoothing parameter. Therefore, in the operation phase, the output of the MMSE beamformer at time-frame k and frequency f is $\mathbf{w}^{\mathtt{k}}(f)^{\mathtt{H}}\mathbf{Y}(f,k)$.

III. Distributed Computation Scheme

In this section, we design a distributed computation scheme of the MMSE beamformer in which the nodes in the WASN are divided into different groups. In our distributed computation scheme, gossip algorithms are used to solve the consensus problem among groups. For different grouping rules, the gossip algorithm with same number of groups may have different convergence speed. In [17], it was suggested that the convergence time of the gossip algorithm depends on the spectral gap of the graph which consists of groups in our case. Taking the network topology into consideration, one can group the nodes with a larger spectral gap so as to speech up the convergence of the gossip algorithm. If the network topology is fixed, one can minimize the convergence time by optimizing the pairwise gossiping probabilities [16]. In practice, we allocate the geographic adjacent nodes into the same group because nodes in the same group need to synchronize the status of private ledgers when gossip algorithms are applied. Apart from that, it is also necessary to consider the size of group since the number of duplication ledgers is increased with the increment of number of nodes in one group, although less groups speed up the convergence of the algorithm.

A. Gossip algorithms

Gossip algorithms are widely used to solve the average consensus problem in decentralized network systems. The randomized gossip algorithm has been used to design a distributed delay-and-sum beamformer without the consideration of transmission failure [13] . They allow nodes exchanging information peer-to-peer and updating the parameter by computing the pairwise average. Eventually, all the nodes in the network agree on the value of the parameter. There are several variants of the gossip algorithms. The main difference between them is the choice of neighbors or routings. We introduce two distributed computation algorithms: the randomized gossip algorithm and greedy gossip algorithm. In the randomized gossip algorithm, the neighbor is chosen uniformly at random from a predefined neighbor set. It has been shown, in [34], that this algorithm converges to a consensus if the graph is strongly connected. Like other greedy algorithms, the greedy gossip algorithm [35] makes an optimal choice among neighbors to achieve a fast convergence. In this paper, the selection criterion of the greedy gossip algorithm is defined as

$$v_2 = \arg\max_{v_2 \in \mathcal{N}_{v_1}} \|\mathbf{x}_{v_1} - \mathbf{x}_{v_2}\|$$

where v_1 is a random chosen group; $\mathcal{N}_{v_1}$ is a predefined neighbor set of Group v_1; v_2 is the chosen neighbor of Group v_1; and $\|\mathbf{x}_{v_1} - \mathbf{x}_{v_2}\| = \sqrt{(\mathbf{x}_{v_1} - \mathbf{x}_{v_2})^{\mathtt{H}}(\mathbf{x}_{v_1} - \mathbf{x}_{v_2})}$ is the Euclidean norm between complex vectors $\mathbf{x}_{v_1}$ and $\mathbf{x}_{v_2}$. It means that the greedy gossip algorithm always chooses the neighbor with the most different value. Comparing with the randomized gossip algorithm, the greedy gossip algorithm accelerates the convergence to a consensus state in the network. However, an additional bandwidth is needed to eavesdrop the information from neighbors. Therefore, both gossip algorithms are investigated in the simulation study.

B. Distributed Computation of MMSE beamformer

Our objective in this subsection is to estimate the $\check{\mathbf{R}}(f,k)$ in Equation (11) distributively but a consensus should be achieved in the network. With the estimation of $\check{\mathbf{R}}(f,k)$ as well as other estimations estimated in the time frame $k-1$ and the optimization phase, it is easy to calculate the optimal beamformer weight in Equation (12) so as to derive the output of MMSE beamformer $\hat{Y}(f,k)$ in Equation (3).

We rewrite Equation (11) as

$$\check{\mathbf{R}}(f,k) = \lambda^{-1}\hat{\mathbf{R}}(\mathbf{f},\mathbf{k}-\mathbf{1})^{-1} - \frac{\lambda^{-2}\mathbf{a}\mathbf{a}^{\mathbf{H}}}{1+\lambda^{-1}b} \tag{13}$$

where

$$\begin{aligned}\mathbf{a} &= \hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k),\\ b &= \mathbf{Y}(f,k)^{\mathrm{H}}\hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k) = \mathbf{Y}(f,k)^{\mathrm{H}}\mathbf{a}.\end{aligned}$$

In the sequel, the gossip algorithm is applied to estimate both $\mathbf{a}$ and b sequentially.

For the v-th group, let $\hat{\mathbf{R}}_v(f,k-1)^{-1}$ be the local estimate of $\hat{\mathbf{R}}(f,k-1)^{-1}$ and $\mathbf{c}_{v,1},\ldots,\mathbf{c}_{v,N}$ be the columns of $\hat{\mathbf{R}}_v(f,k-1)^{-1}$. Let M_v be the set of microphones belonging to the v-th group. Without communication to other groups, we can compute a part of $\mathbf{a}$ in the v-th group by

$$\mathbf{a}^{(v)} = \sum_{i\in M_v} \mathbf{c}_{v,i} Y_i(f,k). \tag{14}$$

Since

$$\mathbf{a} = \sum_{v=1}^{V}\sum_{i\in M_v} \mathbf{c}_{v,i} Y_i(f,k) = \sum_{v=1}^{V} \mathbf{a}^{(v)} = \frac{1}{V}\sum_{v=1}^{V} \tilde{\mathbf{a}}^{(v)}$$

where V is the number of groups and $\tilde{\mathbf{a}}^{(v)} = V\mathbf{a}^{(v)}$, it can be recognized that $\mathbf{a}$ is the arithmetic mean of $\tilde{\mathbf{a}}^{(v)}$. Therefore, $\mathbf{a}$ can be calculated by the gossip algorithm.

Let $a^{(v_i,t)}$ denote the local estimate of $\mathbf{a}$ in Group v_i which are selected to exchange information at the t-th time by the gossip algorithm. The initial value of the local estimate of $\mathbf{a}$ in Group v_i is defined as

$$\mathbf{a}^{(v_i,0)} = \tilde{\mathbf{a}}^{(v_i)}. \tag{15}$$

In one iteration of the gossip algorithm in which Group v_i is selected at the t_i-th time and Group v_j is selected at the t_j-th time, the local estimate of $\mathbf{a}$ in Group v_i and v_j are updated by

$$\mathbf{a}^{(v_i,t_i)} = \mathbf{a}^{(v_j,t_j)} = \frac{1}{2}\left(\mathbf{a}^{(v_i,t_i-1)} + \mathbf{a}^{(v_j,t_j-1)}\right). \tag{16}$$

Let $T_{\mathbf{a},v}$ be the number of times that the v-th Group is selected by the gossip algorithm when $\check{\mathbf{R}}(f,k)$ is estimated. The final local estimate of $\hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k)$ in the v-th Group is

$$\mathbf{a}^{(v,T_{\mathbf{a},v})} = [a_1^{(v,T_{\mathbf{a},v})},\ldots,a_M^{(v,T_{\mathbf{a},v})}].$$

After the estimation work of $\mathbf{a}$, we can estimate b using the same method. Without communication to other groups, we can compute a part of b in the v-th Group by

$$b^{(v)} = \sum_{i\in M_v} Y_i(f,k)^* a_i^{(v,T_{\mathbf{a},v})}. \tag{17}$$

Then,

$$b = \sum_{v=1}^{V}\sum_{i\in M_v} Y_i(f,k)^* a_i^{(v,T_v)} = \frac{1}{V}\sum_{i=1}^{V} V b^{(v)} = \frac{1}{V}\sum_{i=1}^{V}\tilde{b}^{(v)}.$$

It is obvious that b is the arithmetic mean of $\tilde{b}^{(v)}$. Therefore, b can be calculated by the gossip algorithm.

Let $b^{(v_i,t)}$ denote the local estimate of b in Group v_i which are selected to exchange information at the t-th time by the gossip algorithm. The initial value of the local estimate of b in Group v_i is defined as

$$b^{(v_i,0)} = \tilde{b}^{(v_i)}. \tag{18}$$

For the iteration of the gossip algorithm in which Group v_i is selected at the t_i-th time and Group v_j is selected at the t_j-th time, the local estimate of b in Group v_i and v_j are calculated by

$$b^{(v_i,t_i)} = b^{(v_j,t_j)} = \frac{1}{2}\left(b^{(v_i,t_i-1)} + b^{(v_j,t_j-1)}\right). \tag{19}$$

Let $T_{b,v}$ be the number of times that the v-th Group is selected by the gossip algorithm when $\check{\mathbf{R}}(f,k)$ is estimated. The final local estimate of $\mathbf{Y}(f,k)^{\mathrm{H}}\hat{\mathbf{R}}(f,k-1)^{-1}\mathbf{Y}(f,k)$ in the v-th Group is $b^{(v,T_{b,v})}$.

To summarize, in the gossip algorithm, the local estimates of $\mathbf{a}$ and b have exchanged among groups iteratively. The main purpose of the exchange is to calculate the matrix $\check{\mathbf{R}}(f,k)$. After sufficient exchanges, the local estimates of $\check{\mathbf{R}}(f,k)$ converge to the same matrix because $\mathbf{a}$ and b are converge to the same vector and scalar respectively. The algorithm can be summarized in Algorithm 1. Using the local estimate of $\check{\mathbf{R}}(f,k)$, a local beamformer weight vector is derived. The local output of MMSE beamformer in each group is derived using Equation (10), (12) and (3). It is noted that $\mathbf{Y}(f,k)$ is partially unknown for an individual group but its local estimate in the v-th group can be computed by

$$\mathbf{Y}_v(f,k) = [\hat{\mathbf{R}}_v(f,k-1)^{-1}]^{-1}\mathbf{a}^{(v,T_{a,v})}. \tag{20}$$

Algorithm 1 Estimate $\check{\mathbf{R}}(f,k)$ using the gossip algorithm

1: Initialize $\mathbf{a}^{(v,0)}$ for Group v using Equation (14) and (15), where $v=1,\ldots,V$, and let $T=0$.
2: **repeat**
3: Select two groups and update the local estimates of $\mathbf{a}$ using Equation (16).
4: $T=T+1$.
5: **until** $T > T_{\mathbf{a},\max}$ where $T_{\mathbf{a},\max}$ is the hyperparamter denoting the maximum iteration number when $\mathbf{a}$ is estimated.
6: Initialize $b^{(v,0)}$ for Group v using Equation (17) and (18), where $v=1,\ldots,V$, and let $T=0$.
7: **repeat**
8: Select two groups and update the local estimates of b using Equation (19).
9: $T=T+1$.
10: **until** $T > T_{b,\max}$ where $T_{b,\max}$ is the hyperparamter denoting the maximum iteration number when b is estimated.
11: Calculate $\check{\mathbf{R}}_v(f,k)$, the local estimate of $\check{\mathbf{R}}(f,k)$ in the v-th group, by the substitution of $\mathbf{a}^{(v,T_{\mathbf{a},v})}$ and $b^{(v,T_{b,v})}$ into Equation (13).

IV. Blockchain protection

Due to the complexity of computation and network involved in the normal blockchain version, we only use the basic blockchain functions including: adding block, hashing block, block validations and the longest chain rule. Each blockchain has one ledger system to store data. The ledger is an append-only block structure in which the block cannot be removed or modified once it has been added. In order to cooperate with gossip algorithms, there are two ledger systems to separately store the local estimates of both $\mathbf{a}$ and b in one group. When

the network begins to estimate $\check{\mathbf{R}}(f,k)$, the blockchains related to $\check{\mathbf{R}}(f,k-1)$ are removed and new blockchains are initialized. Therefore, the blockchains do not need much memory space.

Fig. 2 demonstrates the ledger structure of the blockchain group in Fig. 1(b) when $\hat{\mathbf{R}}(f,k)^{-1}$ is estimated. Each node in the group has a copy of both Ledger 1 and 2. One block in the blockchain contains index, timestamp, data and hash value. The block except for the genius block also contains a pervious hash value. The data stored in the genius block contains $\hat{\mathbf{R}}_v(f,k-1)^{-1}, \gamma_i(f)$ and $\mathbf{q}_i(f)$ which are the essential data to calculate the local estimates $\hat{\mathbf{R}}_v(f,k)^{-1}$ using Equation (13) and (10). It is noted that, for different groups, $\mathbf{a}^{(v,0)}$ and $b^{(v,0)}$ defined by Equation (15) and (18) should be different such that the hash value for each group cannot be the same. Moreover, the inter-group communication is dominated by gossip algorithms. The selected time for each group in the inter-group communication is at random. Thus, the lengths of Ledger 1 and 2 in each blockchain group could be different. In our paper, we use the fully private blockchain which is open for other groups for reading but the permission to write it belongs to the nodes inside the group itself.

The hash value is generated based on the index, timestamp, data and the previous hash. It can be considered as a fingerprint of the block. When anything in the block is broken, the hash value will be changed. Since the hash value is linked to the next block, the hash values of all the blocks after this block should also be changed. Compared to other healthy chain, it is easy to find out the broken chain or broken block. In the inter-group communication, another groups can easily verify the correctness of the received data using the hash function.

The advantages of using blockchain in the WASN are the following:

- It perfectly resolves the data consensus problem inside the group since data is immutable and tamper-proof in blockchain.
- It accelerates the gossip algorithm since the graph is compressed by the non-overlapping blockchain group.
- Once a communication link is broken, same data can be transmitted from others in the same blockchain group.
- Broken data in one node can be recovered by duplicated data from other nodes in the same blockchain group.

With the increment of number of nodes in the blockchain group, the number of duplication ledgers is increased. It leads to more bandwidth consumption because blockchain technique synchronizes the status of duplication ledgers. Therefore, we use a two-level communication scheme and employ the blockchain technique in small separate groups rather than in the whole WASN.

V. Simulation

In this section, we illustrate the performance of the designed beamformer in a simulated room. Firstly, we consider a $10m \times 10m \times 3m$ square office room with a reverberation time of $T_{60} = 0.2s$. The heights of microphones and sources are 1.5 meters. The horizontal positions of both sources and microphones are shown in Fig. 3. The size of our simulated room is a common specification for a large conference room or a small lecture theater. A similar size of our acoustic system is popular in the literature of studying the distributed beamforming [36]. We construct the blockchain groups according to the set $\{(\texttt{Node 1}, \texttt{Node 2}, \texttt{Node 3}), (\texttt{Node 4}, \texttt{Node 5}, \texttt{Node 6}), (\texttt{Node 7}, \texttt{Node 8}), (\texttt{Node 9}, \texttt{Node 10})\}$. Since signals are generated based on a signal propagation model with a known source location via Equation (1) as in [31], we can randomly select a node to be the reference signal in our model. One microphone of Node 8 is randomly selected to be the reference microphone. Fig. 4 displays the RIR vector for the reference microphone. Furthermore, both source speech and interference speech contain $4s$ voice signals sampled at 16kHz. All signals are transformed in the frequency domain by a 256-tap FIR filter. The over-lapping rate is 50%. In this experiment, the signal-to-interference (SIR) ratio is fixed at -5dB.

Four performance measures are used to evaluate the performance in different cases. Define $\hat{P}_Y(\omega)$ as the spectral power estimate of the source signal; $\hat{P}_{\tilde{Y}}(\omega)$ as the spectral power estimate of the output of beamformer; $\hat{P}_{Y_I}(\omega)$ as the spectral power estimate of the interference speech; $\hat{P}_{\tilde{Y}_I}(\omega)$ as the spectral power estimate of the output of beamformer when the interference speech is active alone; $\hat{P}_{Y_{\texttt{WN}}}(\omega)$ as the spectral power estimate of the white noise; $\hat{P}_{\tilde{Y}_{\texttt{WN}}}(\omega)$ as the spectral power estimate of the output of beamformer when the white noise is active alone. Then, the first performance measure is the normalized distortion which is formulated as

$$\texttt{Distortion} = \frac{1}{\pi}\int_{-\pi}^{\pi} |C_d \hat{P}_{\tilde{Y}}(\omega) - \hat{P}_Y(\omega)| d\omega$$

where C_d is defined as

$$C_d = \frac{\int_{-\pi}^{\pi} \hat{P}_Y(\omega) d\omega}{\int_{-\pi}^{\pi} \hat{P}_{\tilde{Y}}(\omega) d\omega}.$$

The second and third performance measures are the normalized white noise suppression and normalized interference suppression which are respectively defined by

$$\begin{aligned} \texttt{SUPP}_{\texttt{WN}} &= \frac{\int_{-\pi}^{\pi} \hat{P}_{\tilde{Y}_{\texttt{WN}}}(\omega) d\omega}{C_d \int_{-\pi}^{\pi} \hat{P}_{Y_{\texttt{WN}}}(\omega) d\omega}, \\ \texttt{SUPP}_{\texttt{I}} &= \frac{\int_{-\pi}^{\pi} \hat{P}_{\tilde{Y}_{\texttt{I}}}(\omega) d\omega}{C_d \int_{-\pi}^{\pi} \hat{P}_{Y_{\texttt{I}}}(\omega) d\omega}. \end{aligned}$$

The fourth measure is the segmental SNR ratio computed by

$$\texttt{SNR}_{\texttt{seg}} = \frac{1}{K_2}\sum_{k=1}^{K_2} 10\log_{10} \frac{\sum_{f=1}^{F} |Y_i(f,k)|^2}{\sum_{f=1}^{F} |\tilde{Y}(f,k) - Y_i(f,k)|^2}$$

where $Y_i(f,k)$ is the STFT coefficient of the clean speech received by the i-th microphone and F is the total number of frequency bins. The first three methods have been used to measure the performance of centralized MMSE beamformer in [37].

In order to show the importance of blockchain protection, we simulate a poor network environment with heavy packet loss. The rate of transmission failure in the WASN is $1e-4$. If the transmission fails, the receiver will get the null data from

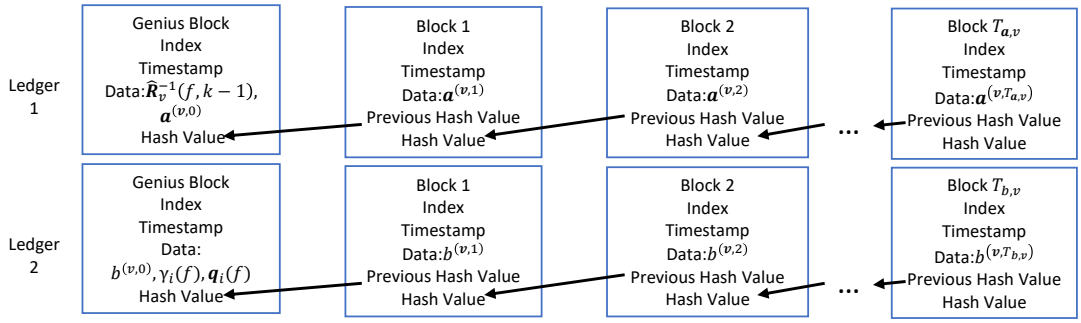


Fig. 2. Designed Data Structure of one blockchain group in WASN

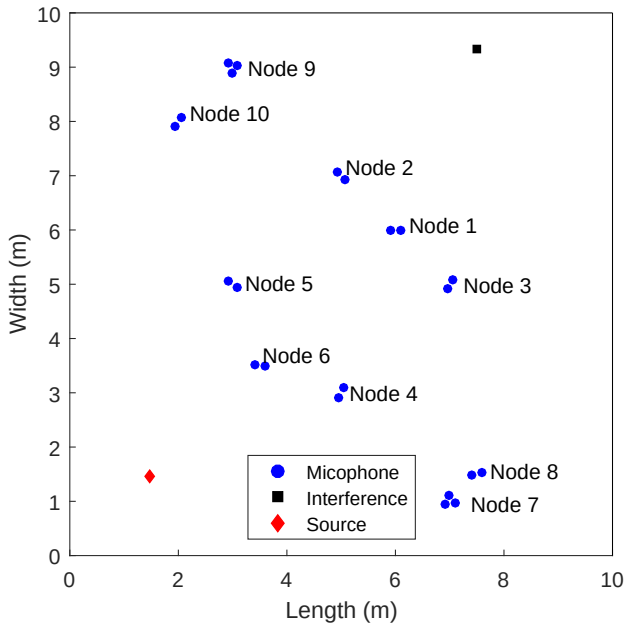


Fig. 3. Locations of 10 nodes (22 microphones), 1 source speech and 1 interference speech in the simulated room

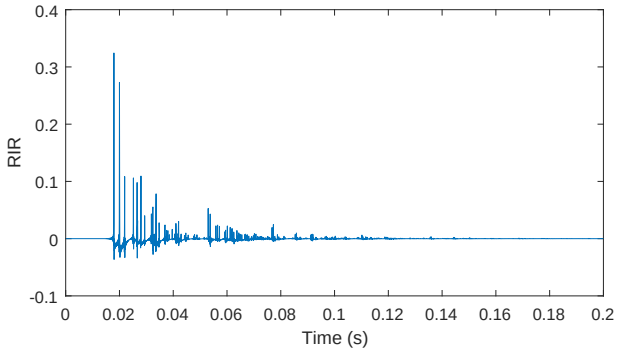


Fig. 4. The room impulse response vector for the reference microphone

the network. We proposed three distributed computational methods below.

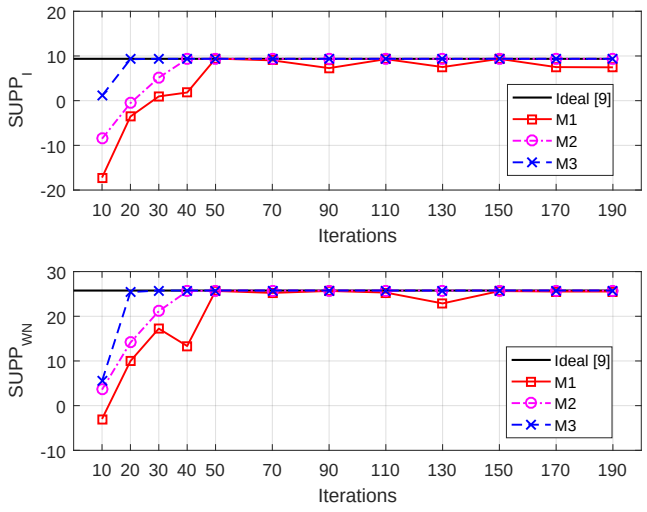
- Method 1 (M1): The MMSE beamformer is distributively computed by the randomized gossip algorithm.
- Method 2 (M2): The MMSE beamformer is distributively computed by the randomized gossip algorithm. Blockchain protection is active.
- Method 3 (M3): The MMSE beamformer is distributively computed by the greedy gossip algorithm. Blockchain protection is active.

For comparison, we calculate the optimal beamformer with all the data readily available in one node, that is, like the centralized counterpart beamformer in which each node would have access to the full set of microphone signals. This is the ideal situation in which all data can be accessed and without any lost of precision due to communication failure. This ideal beamformer has been investigated in [9] and will be used as a benchmark in the following comparison. Moreover, we make a comparison with the distributed delay-and-sum beamformer (DDSB) in [13] as well as the distributed minimum variance distortionless response beamformer (DMVDRB) in [36]. Both DDSB and DMVDRB are calculated by the randomized gossip algorithm.

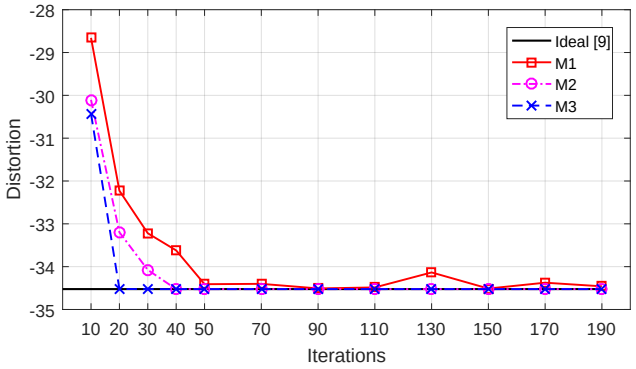
A. Results with a fixed SNR

The input SNR is fixed as -5dB. We investigate the performance of the distributed MMSE beamformer with and without the blockchain protection. Besides that, we compare difference between the randomized gossip algorithm and greedy gossip algorithm when the blockchain protection is active.

Fig. 5 shows the values of interference suppression, white noise suppression and the distortion with the change of iteration numbers in gossip algorithms. As we can see from Fig. 5(a), the beamformer reduces both the white noise and interference speech when the iteration number in gossip algorithms is greater than 30. It is clear that the performance of beamformer is improved over iterations. As we mentioned in Section III, the greedy gossip algorithm used in M3 converges faster because optimal choices among neighbors have been selected. However, the greedy gossip algorithm needs additional resources to eavesdrop the information of neighbors in real time. The red line (M1) denotes the performance of the beamformer which is distributively calculated without blockchain data protection. From Fig. 5(a) and 5(b), we can observe that distributed beamformers with blockchain protection outperform the one without protection. Moreover, the performance of the distributed beamformer is unstable without the blockchain protection. Significant departure on all



(a) Suppression measures of the interference speech (upper) and white noise (lower) with the change of the iteration numbers in gossip algorithms



(b) Speech distortion measure with the change of the iteration numbers in gossip algorithms

Fig. 5. The suppression measures and distortion measure in three methods when the SIR is -5dB and the SNR is -5dB

three performance measures can be observed in M1 even when the iteration number is large.

Fig. 6 depicts the result using M2, namely the beamformered speech (the output of the beamformer), pure speech, noisy speech and the interference speech. It is observed that the distributed beamformer with blockchain protection is able to reduce the noise significantly.

B. Results with different SNRs

In this subsection, the SNR are chosen as -10dB, -5dB, 0dB and 5dB. The iteration number in the randomized gossip algorithm is 110. Fig. 7 shows the segmental SNR versus the input SNR. DMVDRB is calculated by the CbDECM$_1$ algorithm in [36]. DDSB is calculated by the randomized gossip algorithm with clique in [13]. It is observed that the distributed MMSE beamformers outperform both DDSB and DMVDRB in the reverberation environment. Furthermore, the distributed MMSE beamformer using M2 can achieve the performance of the ideal MMSE beamformer. Note that the performances of both the distributed beamformer using M2

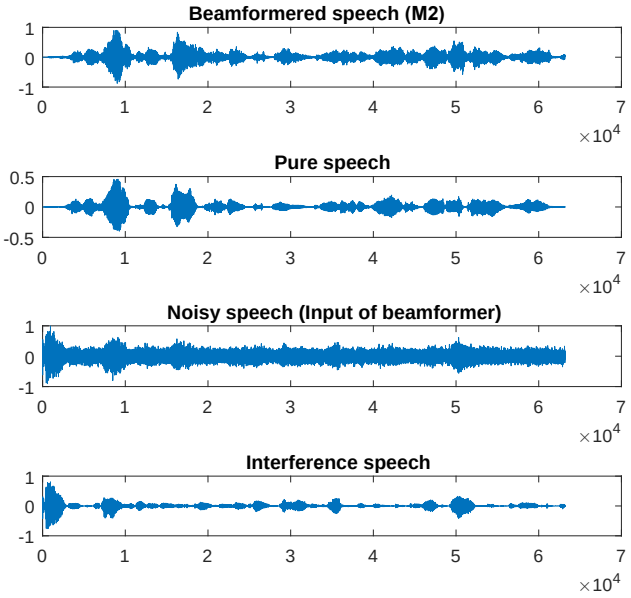


Fig. 6. Beamforming performance of the distributed MMSE beamformer with the blockchain protection (M2)

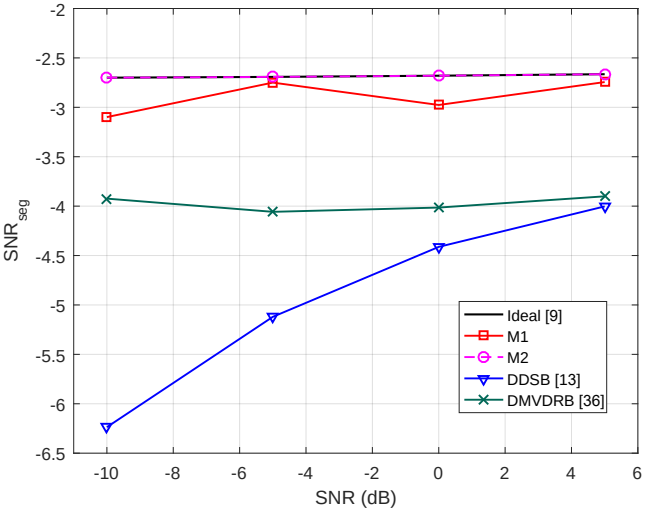


Fig. 7. Segmental SNR with the input SNR chosen as -10dB, -5dB, 0dB and 5dB

and ideal beamformer are the same and therefore these two plots are overlapped in the Figure. Comparing with M2, the performance of M1 is obviously poorer due to the lack of blockchain protection when transmission errors exist.

C. Error analysis

Since $\mathbf{a}$ and b are transmitted to calculate $\check{\mathbf{R}}(f,k)^{-1}$, we define the estimated error rate of $\check{\mathbf{R}}(f,k)^{-1}$, in the v-th blockchain group, as

$$\mathtt{Err}_v(k) = \frac{1}{F}\sum_{f=1}^{F} \frac{\|\check{\mathbf{R}}_v(f,k)^{-1} - \check{\mathbf{R}}_i(f,k)^{-1}\|_{\mathtt{F}}}{\|\check{\mathbf{R}}_i(f,k)^{-1}\|_{\mathtt{F}}} \qquad (21)$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius Norm of the matrix. $\check{\mathbf{R}}_v(f,k)^{-1}$ and $\check{\mathbf{R}}_i(f,k)^{-1}$ are the estimates of $\check{\mathbf{R}}(f,k)^{-1}$ from the distributed beamformer and ideal beamformer. In this case, we focus on the error rate of the third blockchain group when SNR is equal to -5dB. Fig. 8 displays the comparison of the cumulative mean of the estimated error rate of $\check{\mathbf{R}}(f,k)^{-1}$ in both M1 and M2. It is observed that the cumulative mean of $\mathtt{Err}_3$ in M2 roughly behaves like a horizontal line. It means, in M2, that the randomized gossip algorithm attains a stable state and the estimate of $\check{\mathbf{R}}(f,k)^{-1}$ converges to the estimate of $\check{\mathbf{R}}(f,k)^{-1}$ from the ideal beamformer. Without the blockchain protection, the cumulative means of $\mathtt{Err}_v$ in M1 are larger than the counterpart in M2. Apart from that, in the result for M1, there are several jumps which are caused by the transmission failure in the WASN.

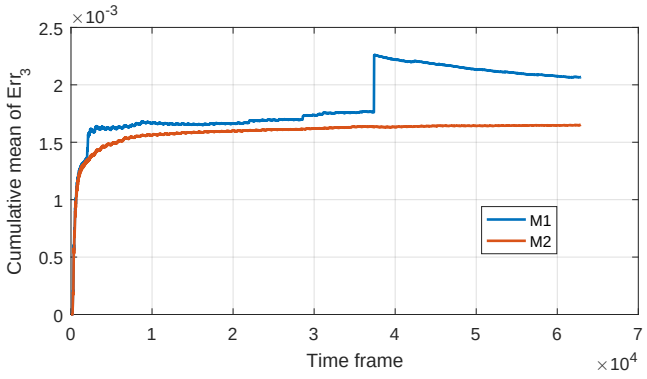


Fig. 8. Estimation of error rate of $\check{\mathbf{R}}(f,k)^{-1}$ on the third blockchain group for M1 and M2 at SNR=-5dB.

VI. Discussion

From the comparison results depicted in Fig. 5, we observe that the performance of the distributed MMSE beamformer without blockchain protection (M1) is worst and unstable when it is put under a poor network environment with heavy packet loss. The designed beamformer drifts away from the optimal performance from time to time, and it takes a while for it to converge again. On the other hand, the proposed distributed MMSE beamformer with blockchain protection (M2) can approach the performance of the ideal beamformer and stay at the optimal level. Moreover, it we elaborate on the iterative technique further by employing the greedy gossip algorithm instead of the randomized gossip algorithm, the distributed MMSE beamformer with blockchain protection (M3) can converge even faster, with the trade-off for using additional resources to make the optimal choice among neighbors in each iteration. The signal outputs from the distributed MMSE beamformer with blockchain protection (M2) is shown in Fig. 7. From the beamformed speech output, it can be seen that the white noise and interference speech have been by and large filtered out, leaving a good estimate of the required speech signal.

Fig. 6 displays the segmental SNR performance of output signals for a range of input SNRs. Compared with two existing implementations of distributed beamforming in the literature, namely DDSB [13] and DMVDRB [36], the distributed MMSE beamformers (M1 and M2) outperforms them in the indoor environment with reverberation. Furthermore, the proposed distributed MMSE beamformer with blockchain protection (M2) is the only one that can achieve the performance of the ideal beamformer for the whole range of SNRs in the study. Without blockchain protection, the performance of M1 often deviates from the optimal performance.

To further understand the cause of the deviation, we conduct an error analysis and find that it is mainly due to errors in the computation of $\check{\mathbf{R}}(f,k)^{-1}$. In Fig. 8. An error measure is introduced in Equation (21) to quantify the effect. It can be seen that faulty data causes the matrix to drift away from the correct value and hence induce performance degradation in the designed beamformer. It is also evident that this problem can be avoided with blockchain protection, which reduces the computational error caused by poor transmission.

VII. Conclusion and Further Work

In this paper, we propose three distributed MMSE beamformers using gossip algorithms. Our proposed distributed MMSE beamformers outperform the distributed delay-and-sum beamformer and MVDR beamformer in a simulated reverberation environment. Based on the blockchain technique, a data protection scheme is also proposed to avoid faulty data transmissions. To illustrate the effectiveness of the proposed beamformer with the blockchain protection, we simulated a typical scenario in a square office room with reverberation. The experimental results show that blockchain data protection is able to secure the quality of the output signals from a distributed beamformer in a relatively poor network environment. In addition, we showed that greedy algorithm can perform better than the randomized gossip algorithm if additional sources are used to receive information from all neighbors. As a future extension, it is of interest to study and optimize the proposed beamformer with voice control accuracy as the performance criteria in smart systems.

Acknowledgment

This work is supported by RGC Grant PolyU 152245/18E and PolyU Grant 4-ZZGS. The second author is also supported by the AMSS-PolyU JLab Postdoctoral fellowship scheme.

References

[1] S. Haller, S. Karnouskos, and C. Schroth, "The internet of things in an enterprise context," in *Future Internet Symposium*. Springer, 2008, pp. 14–28.

[2] R. V. Cox, C. A. Kamm, L. R. Rabiner, J. Schroeter, and J. G. Wilpon, "Speech and language processing for next-millennium communications services," *Proceedings of the IEEE*, vol. 88, no. 8, pp. 1314–1337, 2000.

[3] X. Han and M. Rashid, "Gesture and voice control of internet of things," in *2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2016, pp. 1791–1795.

[4] Y. Meng, Z. Wang, W. Zhang, P. Wu, H. Zhu, X. Liang, and Y. Liu, "Wivo: Enhancing the security of voice control system via wireless signal in iot environment," in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2018, pp. 81–90.

[5] M. Vacher, B. Lecouteux, and F. Portet, "Recognition of voice commands by multisource asr and noise cancellation in a smart home environment," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. IEEE, 2012, pp. 1663–1667.

[6] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 25, no. 4, pp. 692–730, 2017.

[7] S. Nordholm, I. Claesson, and M. Dahl, "Adaptive microphone array employing calibration signals: an analytical evaluation," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 3, pp. 241–252, 1999.

[8] S. Markovich-Golan, A. Bertrand, M. Moonen, and S. Gannot, "Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks," *Signal Processing*, vol. 107, pp. 4–20, 2015.

[9] S. Nordholm, I. Claesson, and N. Grbić, "Optimal and adaptive microphone arrays for speech input in automobiles," in *Microphone Arrays*. Springer, 2001, pp. 307–329.

[10] J. Zhang, S. P. Chepuri, R. C. Hendriks, and R. Heusdens, "Microphone subset selection for mvdr beamformer based noise reduction," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 550–563, 2017.

[11] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *2011 18th IEEE symposium on communications and vehicular technology in the Benelux (SCVT)*. IEEE, 2011, pp. 1–6.

[12] A. Bertrand and M. Moonen, "Distributed node-specific lcmv beamforming in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 233–246, 2011.

[13] Y. Zeng, R. C. Hendriks, and R. Heusdens, "Clique-based distributed beamforming for speech enhancement in wireless sensor networks," in *21st European Signal Processing Conference (EUSIPCO 2013)*. IEEE, 2013, pp. 1–5.

[14] R. Heusdens, G. Zhang, R. C. Hendriks, Y. Zeng, and W. B. Kleijn, "Distributed mvdr beamforming for (wireless) microphone networks using message passing," in *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*. VDE, 2012, pp. 1–4.

[15] S. Doclo, M. Moonen, T. Van den Bogaert, and J. Wouters, "Reduced-bandwidth and distributed mwf-based noise reduction algorithms for binaural hearing aids," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 1, pp. 38–51, 2009.

[16] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[17] A. G. Dimakis, S. Kar, J. M. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[18] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.* IEEE, 2003, pp. 482–491.

[19] H. Zheng and J. Boyce, "An improved udp protocol for video transmission over internet-to-wireless networks," *IEEE Transactions on Multimedia*, vol. 3, no. 3, pp. 356–365, 2001.

[20] J. Postel, "User datagram protocol," *Request for Comments, RFC 768, ISI*, 1980.

[21] P. Koopman and T. Chakravarty, "Cyclic redundancy code (crc) polynomial selection for embedded networks," in *International Conference on Dependable Systems and Networks, 2004*. IEEE, 2004, pp. 145–154.

[22] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "Spins: Security protocols for sensor networks," *Wireless networks*, vol. 8, no. 5, pp. 521–534, 2002.

[23] H. Chan, A. Perrig, B. Przydatek, and D. Song, "Sia: Secure information aggregation in sensor networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 69–102, 2007.

[24] I. Kamel and H. Juma, "A lightweight data integrity scheme for sensor networks," *Sensors*, vol. 11, no. 4, pp. 4118–4136, 2011.

[25] F. Lalem, M. Alshaikh, A. Bounceur, R. Euler, L. Laouamer, L. Nana, and A. Pascu, "Data authenticity and integrity in wireless sensor networks based on a watermarking approach," in *The Twenty-Ninth International Flairs Conference*, 2016.

[26] T. M. Fernández-Caramés and P. Fraga-Lamas, "A review on the use of blockchain for the internet of things," *IEEE Access*, vol. 6, pp. 32 979–33 001, 2018.

[27] T. Braun, T. Voigt, and A. Dunkels, "Tcp support for sensor networks," in *2007 Fourth Annual Conference on Wireless on Demand Network Systems and Services*. IEEE, 2007, pp. 162–169.

[28] M. Pilkington, "11 blockchain technology: principles and applications," *Research handbook on digital transformations*, vol. 225, 2016.

[29] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman, "Blockchain technology: Beyond bitcoin," *APPLIED INNOVATION*, p. 6, 2016.

[30] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for iot," in *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*. ACM, 2017, pp. 173–178.

[31] Z. Li, K. F. C. Yiu, and S. Nordholm, "On the indoor beamformer design with reverberation," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 22, no. 8, pp. 1225–1235, 2014.

[32] N. Grbic, S. Nordholm, J. Nordberg, and I. Claesson, "A new pilot-signal based space-time adaptive algorithm," in *In Proc. of IEEE International Conference on Communications, ICT 2001*, 2001.

[33] M. S. Bartlett, "An inverse matrix adjustment arising in discriminant analysis," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 107–111, 1951.

[34] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2508–2530, 2006.

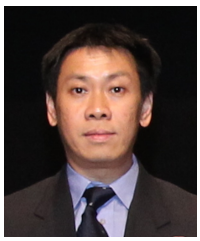
[35] D. Ustebay, B. N. Oreshkin, M. J. Coates, and M. G. Rabbat, "Greedy gossip with eavesdropping," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3765–3776, 2010.

[36] Y. Zeng and R. C. Hendriks, "Distributed estimation of the inverse of the correlation matrix for privacy preserving beamforming," *Signal Processing*, vol. 107, pp. 109–122, 2015.

[37] K. F. C. Yiu, N. Grbic, K.-L. Teo, and S. Nordholm, "A new design method for broadband microphone arrays for speech input in automobiles," *IEEE Signal Processing Letters*, vol. 9, no. 7, pp. 222–224, 2002.

Qingzheng Wang received his B.Sc. degree in computer science and technology from Henan University, Kaifeng, China. He received his M.Sc. degree in operational research and risk analysis and M.Phil. degree in applied statistics from the Hong Kong Polytechnic University, Kowloon, Hong Kong, in 2015 and 2018, respectively. He is currently pursuing his Ph.D. degree in the Department of Applied Mathematics at the Hong Kong Polytechnic University. His research interests include the optimization of the distributed acoustic beamforming system, blind source separation, multi-dimensional constellation design and Bayesian time series analysis.

Shan Guo received the B.Sc. degree in Mathematics from Zhejiang University, Hangzhou, China, in 2012 and the Ph.D. degree in operation research and cybernetics from School of Mathematical Sciences, Zhejiang University, Hangzhou, China, in 2018. Currently, she is a postdoctoral fellow in the Department of Applied Mathematics, the Hong Kong Polytechnic University. Her research interests include optimal control, computational optimization and computational inverse problems. Her current work also includes optimizing applications in areas such as communications and image processing.



Ka-fai Cedric Yiu received his M.Sc. from University of Dundee and University of London, and D.Phil. from University of Oxford. He had worked closely with the industry on different projects in University of Oxford and University College of London. He started his lecturing career in the University of Hong Kong. He is currently working in the Hong Kong Polytechnic University. He has served on the program committee and the organizing committee of a number of conferences and has organized a number of special sessions in conferences. He has published over 100 journal publications and given over 30 conference presentations. He holds two U.S. patents in signal processing. He received the third prize of Chongqing Natural Science Foundation Award in 2014. He also received the Donald Julius Groen Prize back in 2002. He is currently working on several research projects related to optimization and data analysis, signal processing system, risk management and high frequency trading. His current research interests include optimization and optimal control, signal processing and financial risk management.