An Analysis of Performance Improvement by a Helper for Wireless Sensor Networks

Wensheng Lin* and Tad Matsumoto*

*Japan Advanced Institute of Science and Technology (JAIST)
Email: {linwest, matumoto}@jaist.ac.jp
†Centre for Wireless Communications, University of Oulu, Finland
Email: tadashi.matsumoto@oulu.fi

Abstract—Wireless sensor networks (WSNs) are the core component in the big data era. Due to the unreliable transmission environment, it is significantly useful to introduce a helper to refine the system performance. To begin with, we formulate the system model of WSNs as a problem of multiterminal source coding. Subsequently, we propose a 3D distributed compress-bin scheme and derive a corresponding inner bound by analyzing the expected rate-distortion. Finally, we investigate the performance improvement of a helper by comparing the derived inner bound with the Berger-Tung inner bound and through simulation. Both the theoretical bounds and simulation results indicate that a helper can obviously improve the system performance.

Index Terms—Wireless sensor networks, multiterminal source coding, side information, rate-distortion.

I. INTRODUCTION

With the dawn of big data era, it is of great importance that multiple correlated source information are exchanged via lossy communication networks such as wireless sensor networks (WSNs), where full recovery of the sources is not necessarily required. Recently, helper is increasingly introduced into various communication systems, for the purpose of more reliable transmissions [1]–[4]. For instance, Fig. 1 depicts a scenario of WSNs with a helper. Initially, sensors collect information from the observation targets from diverse positions. Therefore, there are observation errors that already exist in the data collected by the sensors. Then, sensors send observations to fusion center through wireless channels. Simultaneously, a helper forms the helper information by processing the observations monitored from sensors by error-free wired or wireless connections. Subsequently, fusion center combines all the observations and makes final estimation with the information provided by the helper. Such significant scenario with the assistance of a helper has stimulated our interests in the analysis of performance improvement provided by a helper for WSNs.

Essentially, the WSNs can be formulated as a problem of multiterminal source coding, which is a classical system model for theoretical analysis. In-depth investigations on multiterminal source coding have been made during the past several decades [5]–[7]. The theoretical basis of distributed lossless source coding is established by Slepian and Wolf [8], who investigate the admissible rate region for discrete memoryless source (DMS). As for lossy case, Berger [9] and Tung [10] finish the pioneering work for general distributed lossy compression system, in which the recovered data are subject to distortion to some extent. Besides, Wyner and Ziv [11] make their contribution to lossy compression with side information, which can be also viewed as a special case for multiterminal source coding when there is no rate limit on one of the source links.

So far, there are many researchers focus on decreasing the distortion in WSNs. Sartipi [12] presents a lossy distributed compression scheme and analyze the corresponding rate-distortion limit for WSNs. Alsheikh et al. [13] develop a data compression algorithm which exploit the spatio-temporal correlations among data, so as to save power and guarantee distortion simultaneously. Xiao et al. [14] raise a policy to minimize the long-term discounted distortion in the fusion center of WSNs. Leinonen et al. [15] propose a compressed sensing scheme for the trade-off between the signal reconstruction distortion and the average encoding rate. Nevertheless, few researches have paid attention to the analysis of performance improvement by introducing a helper. Hence, investigating the effect of a helper for WSNs is a very interesting topic. By solving this problem, it will also provide us a solution for the trade-off among distortions, source rate and helper rate. Motivated by the background described above, in this paper, we investigate the performance improvement provided by a helper for WSNs.
The contributions of this paper are summarized as follows:

- This paper proposes a novel coding scheme of distributed lossy compression for WSNs with a helper, which exploits the joint typicality to minimize the distortion and/or reduce transmission rate.
- Moreover, we derive an inner bound on rate-distortion region by the analysis of expected distortion of the proposed 3D distributed compress-bin scheme.
- Finally, we illustrate the theoretical bound and design a simulation to investigate the performance improvement provided by a helper. Both the theoretical and simulation results prove that a helper can obviously promote the system performance of WSNs.

The rest of this paper is organized as follows. Section II formulates the theoretical model of WSNs with a helper. The 3D distributed compress-bin scheme and its theoretical inner bound are presented in Section III. Section IV evaluates the performance improvement by introducing a helper through the comparison of theoretical bounds and simulations. Finally, we conclude this work in Section V.

II. PROBLEM STATEMENT

![Fig. 2. The simplest system model of WSNs with a helper.](image)

For the purpose of theoretical analysis, we formulate the WSNs with a helper as a problem of multiterminal source coding with a helper, which is illustrated in Fig. 2. \(X_1\), \(X_2\), and \(Y\) are three memoryless sources, which produce independent and identically distributed (i.i.d.) sequences \(x_1^n = \{x_1(i)\}_{i=1}^n\), \(x_2^n = \{x_2(i)\}_{i=1}^n\) and \(y^n = \{y(i)\}_{i=1}^n\) by taking values from three finite alphabets \(\mathcal{X}_1\), \(\mathcal{X}_2\), and \(\mathcal{Y}\) for each time slot, respectively. Encoder \(j\) observes the target sequence \(x_j^n\) for \(j = 1, 2\), while encoder \(H\) observes the side information sequence \(y^n\). Due to some strict situation in practice, such as encoders are far away from each other, three encoders have to independently encode their own observations. Three observed sequences \(x_1^n\), \(x_2^n\), and \(y^n\) are encoded at rates \(R_1\), \(R_2\), and \(R_H\), separately, by assigning an index to each sequence according to the following mapping rules:

\[
\varphi_j : \mathcal{X}_j^n \mapsto \mathcal{M}_j = \{1, 2, \ldots, 2^{nR_j}\}, \text{ for } j = 1, 2, \\
\varphi_H : \mathcal{Y}^n \mapsto \mathcal{M}_H = \{1, 2, \ldots, 2^{nR_H}\}.
\]

Then, the encoder outputs \(\varphi_1(x_1^n)\), \(\varphi_2(x_2^n)\) and \(\varphi_H(y^n)\) are transmitted to a common receiver. In contrast to distributed compression in encoder, decoder can jointly construct the estimates \(\hat{x}_1^n\) and \(\hat{x}_2^n\) from indices \(\varphi_1(x_1^n)\) and \(\varphi_2(x_2^n)\) by utilizing the side information \(\varphi_H(y^n)\).

The reconstruction process can be implemented by the mappings as:

\[
\psi_j : \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_H \mapsto \mathcal{X}_j^n, \text{ for } j = 1, 2.
\]

The distortion measure \(d_j : \mathcal{X}_j \times \mathcal{X}_j \mapsto [0, \infty)\) is defined to describe the degree of distortion between \(x_j\) and \(\hat{x}_j\), for \(j = 1, 2\). Then, we define the average distortion between the sequences \(x_j^n\) and \(\hat{x}_j^n\) as

\[
d_j(x_j^n, \hat{x}_j^n) = \frac{1}{n} \sum_{i=1}^{n} d_j(x_j(i), \hat{x}_j(i)), j = 1, 2.
\]

For given distortion requirements \((D_1, D_2)\), the rate-distortion region \(\mathcal{R}(D_1, D_2)\), consisting of all achievable rate triples of \((R_1, R_2, R_H)\), is defined as

\[
\mathcal{R}(D_1, D_2) = \{(R_1, R_2, R_H) : (R_1, R_2, R_H) \text{ is admissible and } \lim_{n \to \infty} E(d_j(x_j^n, \hat{x}_j^n)) \leq D_j + \epsilon, \text{ for } j = 1, 2, \text{ and any } \epsilon > 0 \}.
\]

The theoretical goal is to determine the rate-distortion region \(\mathcal{R}(D_1, D_2)\). By utilizing the theoretical result, it is effective to analyze the performance improvement by introducing a helper.

III. ANALYSIS OF PERFORMANCE IMPROVEMENT

In this section, we first propose an achievable source coding scheme, i.e., 3D distributed compress-bin scheme. Then, we derive an inner bound of rate-distortion region based on the 3D distributed compress-bin scheme.

A. 3D Distributed Compress-bin Scheme

As depicted in Fig. 3, the 3D distributed compress-bin scheme uses joint typicality in random binning at each encoder, while it uses joint typicality of decoding from the 3D bin and symbol-by-symbol reconstructions at the decoder. This 3D random binning coding scheme consists of three key techniques—codebook generation, encoding and decoding. In the following, we introduce the three key techniques one by one, with the assumption that \(\epsilon > \epsilon' > \epsilon'' > \epsilon'''\).

**Codebook generation.** Initially, fix a conditional pmf \(p(u_1|x_1)p(u_2|x_2)\) for two sources. Then, fix a conditional pmf \(p(v|y)\) and let \(p(v) = \sum_y p(y)p(v|y)\) for
the helper. Additionally, fix two functions \( \tilde{x}_1(u_1, u_2, v) \) and \( \tilde{x}_2(u_1, u_2, v) \) for the joint decoder such that \( E(d_j(X_j, X_j)) \leq D_j/(1 + \epsilon), \ j = 1, 2 \). Let \( R_1 \geq R_1, R_2 \geq R_2 \); subsequently, we can assume without loss of generality that \( nR_1, nR_2, nR_1, nR_2 \) and \( nR_H \) are integers if \( n \) is sufficiently large. For \( j = 1, 2 \), randomly and independently generate \( 2^{nR_j} \) sequences \( u_j^1(l_j) \sim \prod_{i=1}^n p_{v_j}(u_{ji}), l_j \in [1 : 2^{nR_j}] \). Similarly, randomly and independently generate \( 2^{nR_H} \) sequences \( v^*(m_H) \sim \prod_{i=1}^n p_{v}(v_i), m_H \in [1 : 2^{nR_H}] \). Partition the set of indices \( l_j \in [1 : 2^{nR_j}] \) into equal-size bins \( B_j(m_j) = [(m_j - 1)2^{nR_j}/R_j + 1 : m_j 2^{nR_j}/R_j], m_j \in [1 : 2^{nR_j}] \). The codebook is revealed to the encoders and the decoder.

**Encoding.** Upon observing \( x_j^n \), encoder \( j = 1, 2 \) finds an index \( l_j \in [1 : 2^{nR_j}] \) such that \( (u_j^1(l_j), x_j^n) \in T_j^{(n)}(U_j, X_j) \). If there is more than one such index \( l_j \), then encoder \( j \) selects one of them uniformly at random. If there is no such index \( l_j \), then encoder \( j \) selects an index from \( [1 : 2^{nR_j}] \) uniformly at random. Encoder \( j = 1, 2 \) sends the index \( m_j \) such that \( l_j \in B_j(m_j) \). Similarly, encoder \( H \) finds an index \( m_H \) such that \( (v^*(m_H), y^n) \in T_H^{(n)}(V, Y) \). If there is more than one such index, then encoder \( H \) selects one of them uniformly at random. If there is no such index, then encoder \( H \) selects an index from \( [1 : 2^{nR_H}] \) uniformly at random. Encoder \( H \) then sends the index \( m_H \) to the decoder.

**Decoding.** The decoder finds the unique index pair \( (\tilde{l}_1, \tilde{l}_2) \in B_1(m_1) \times B_2(m_2) \) such that \( (u_1^1(\tilde{l}_1), u_2^1(\tilde{l}_2), v^*(m_H)) \in T_\tilde{H}^{(n)}(U_1, U_2, V) \). If there is such a unique index pair \( (\tilde{l}_1, \tilde{l}_2) \), then the reconstructions are computed as \( \hat{x}_1(\tilde{u}_1, \tilde{l}_1), u_2(\tilde{l}_2, v_1) \) and \( \hat{x}_2(u_1, \tilde{l}_1), u_2(\tilde{l}_2, v_1) \) for \( i \in [1 : n] \); otherwise, \( \hat{x}_1^n \) and \( \hat{x}_2^n \) are set to arbitrary sequences in \( X_1^n \) and \( X_2^n \), respectively.

### B. Analysis of Expected Rate-Distortion

Let \( (L_1, L_2, M_H) \) denote the triplet of indices for the chosen \( (U_1^n, U_2^n, V^n) \) triplet, \( (M_1, M_2) \) be the pair of corresponding bin indices, and \( (L_1, L_2) \) be the pair of decoded indices. Define the “error” event

\[
E = \{ U_1^n(\hat{L}_1), U_2^n(\hat{L}_2), V^n(M_H), X_1^n, X_2^n, Y^n \}
\]

and consider the following events:

\[
\begin{align*}
E_1 &= \{ (V^n(m_H), Y^n) \notin T^{(n)}_e \} \\
E_2 &= \{ (U_j^n(l_j), X_j^n) \notin T^{(n)}_j \} \\
E_3 &= \{ (U_j^n(l_j), X_j^n, Y^n) \notin T^{(n)}_j \} \\
E_4 &= \{ (U_j^n(l_j), X_j^n, V^n(M_H), Y^n) \notin T^{(n)}_e \} \
E_5 &= \{ (U_j^n(l_j), X_j^n, V^n(M_H), Y^n) \notin T^{(n)}_e \} \\
E_6 &= \{ (U_j^n(l_1), U_j^n(l_2), V^n(M_H)) \notin \mathcal{T}^{(n)}_e \}
\end{align*}
\]

where \( E_1 \) and \( E_2 \) represent encoding error events in encoder \( H \) and encoder \( j \) for \( j = 1, 2 \), respectively; \( E_5 \) occurs if joint typicality decoding fails, and \( E_6 \) means that there are more than one decoding result and hence a decoding error event occurs. Since the “error” event occurs only if \( (U_j^n(L_1), U_j^n(L_2), V^n(M_H), X_1^n, X_2^n, Y^n) \notin \mathcal{T}^{(n)}_e \) or \( (L_1, L_2) \neq (L_1, L_2) \), by the union of the events bound,

\[
P(E) \leq P(E_1) + P(E_2) + P(E_3 \cap E_4) + P(E_5 \cap E_4) + P(E_6)
\]

We bound each term as follows. First, by the covering lemma, \( P(E_1) \) tends to zero as \( n \to \infty \) if

\[
R_H > I(Y; V) + \delta(e'''), \quad (2)
\]

and \( P(E_2) \) tends to zero as \( n \to \infty \) if

\[
\tilde{R}_j > I(U_j; X_j) + \delta(e'''), \quad (3)
\]

Since \( P(E_3) = \{ (U_j^n(L_j), X_j^n) \in \mathcal{T}^{(n)}_j \} \), \( Y^n|U_j^n(L_j) = u_j^n, X_j^n = x_j^n \sim \prod_{i=1}^n p_{y|x}(y_i|x_i) \) for \( j = 1, 2 \), and \( e'''>e''' \), by the conditional typicality lemma, \( P(E_3 \cap E_4) \) tends to zero as \( n \to \infty \).

To bound \( P(E_5 \cap E_4) \), for \( j = 1, 2 \), let \( (u_j^n, x_j^n, y^n) \in \mathcal{T}^{(n)}_j(U_j, X_j, Y) \), and consider

\[
P(V^n(M_H) = v^n|U_j^n(L_j)) = u_j^n, X_j^n = x_j^n, Y^n = y^n
\]

\[
P(V^n(M_H) = v^n|Y^n = y^n)
\]

\[
p(v^n|y^n)
\]

(4)
First, note that by the covering lemma, \( P\{V^n(M_H) \in \mathcal{T}_e^n(V^n) | Y^n = y^n\} \) converges to 1 as \( n \to \infty \), that is, \( p(v^n | y^n)\) satisfies the first condition in the Markov lemma. Then, similar to the proof of the Berger-Tung inner bound, through Lemma 12.3 in [16] with \( U_2 \leftarrow V, X_2 \leftarrow Y, p(v^n | y^n)\) also satisfies the second condition in the Markov lemma. Hence, from the Markov lemma with \( Z \leftarrow V, Y \leftarrow Y, \) and \( X \leftarrow (U_j, X_j)\), we have

\[
\lim_{n \to \infty} P\{\{u^n_j, x^n_j, y^n\} \in \mathcal{T}_e^n(U_j^n(L_j)) = u^n_j, X^n_j = x^n_j, Y^n = y^n\} = 1, \quad (5)
\]

if \( (u^n_j, x^n_j, y^n) \in \mathcal{T}_e^n(U_j, X_j, Y) \) and \( \epsilon'' < \epsilon' \) is sufficiently small. Therefore, \( P(\mathcal{E}_3^n \cap \mathcal{E}_4) \) tends to zero as \( n \to \infty \).

Since \( P(\mathcal{E}_3^n) = \{(U^n_1(L_1), X^n_1, V^n(M_H), Y^n) \in \mathcal{T}_e^n \cap \{(U^n_2(L_2), X^n_2, V^n(M_H), Y^n) \in \mathcal{T}_e^n\} \). Now, we bound \( P(\mathcal{E}_3^n \cap \mathcal{E}_5) \) using a similar method of bounding \( P(\mathcal{E}_3^n \cap \mathcal{E}_4) \). Let \( (u^n_1, x^n_1, v^n, y^n) \in \mathcal{T}_e^n \) and consider

\[
P\{(U^n_2(L_2), X^n_2) = (u^n_2, x^n_2) | U^n_1(L_1) = u^n_1, X^n_1 = x^n_1, V^n(M_H) = v^n, Y^n = y^n\}
\]

\[
= P\{(U^n_2(L_2), X^n_2) = (u^n_2, x^n_2) | V^n(M_H) = v^n, Y^n = y^n\}
\]

\[
= p(u^n_2, x^n_2 | v^n, y^n). \quad (6)
\]

First, note that by the covering lemma, \( P\{(U^n_2(L_2), X^n_2) \in \mathcal{T}_e^n(U_2, X_2 | v^n, y^n) | V^n(M_H) = v^n, Y^n = y^n\} \) converges to 1 as \( n \to \infty \), that is, \( p(u^n_2, x^n_2 | v^n, y^n)\) satisfies the first condition in the Markov lemma. Then, through Lemma 12.3 in [16] with \( U_2 \leftarrow (U_2, X_2), X_2 \leftarrow (V, Y), p(u^n_2, x^n_2 | v^n, y^n)\) also satisfies the second condition in the Markov lemma. Hence, from the Markov lemma with \( Z \leftarrow (U_2, X_2), Y \leftarrow (V, Y), \) and \( X \leftarrow (U_1, X_1)\), we have

\[
\lim_{n \to \infty} P\{(u^n_1, x^n_1, v^n, y^n) \in \mathcal{T}_e^n(U_1, X_1, V, Y) | U^n_1(L_1) = u^n_1, X^n_1 = x^n_1, V^n(M_H) = v^n, Y^n = y^n\}
\]

\[
= 1, \quad (7)
\]

if \( (u^n_1, x^n_1, v^n, y^n) \in \mathcal{T}_e^n(U_1, X_1, V, Y) \) and \( \epsilon' < \epsilon \) is sufficiently small. Therefore, \( P(\mathcal{E}_3^n \cap \mathcal{E}_5) \) tends to zero as \( n \to \infty \).

Following a similar argument as Lemma 11.1 in [16] in the proof of the Wyner-Ziv theorem, we have

\[
P(\mathcal{E}_6^n) \leq P\{(\tilde{U}_1^n, \tilde{U}_2^n) \in \mathcal{T}_e^n(V^n(M_H)) \in \mathcal{T}_e^n | (\tilde{l}_1, \tilde{l}_2) \in B_1(1) \times B_2(1)\}. \quad (8)
\]

To bound \( P(\mathcal{E}_6^n) \), we introduce the following two lemmas:

**Lemma 1** (joint typicality lemma for a triplet of random variables): Let \( (U_1, U_2, V) \sim p(u_1, u_2, v) \). If \( \tilde{v} \) is an arbitrary sequence and \( \tilde{U}_n \sim \prod_{i=1}^n p_V(\tilde{u}_{ji}) \) for \( j = 1, 2 \), then

\[
P\{\tilde{U}_1, \tilde{U}_2, \tilde{v} \in \mathcal{T}_e^n(U_1, U_2, V) \}
\]

\[
\leq 2^{-n(H(U_1, U_2, V) - \delta(\epsilon))} \quad (9)
\]

**Lemma 2** (mutual packing lemma for a triplet of random variables): Let \( (U_1, U_2, V) \sim p(u_1, u_2, v) \). Let \( U^n_j(l_j) \sim \prod_{i=1}^n p_{U_j}(u_{ji}), l_j \in L_j = [1 : 2^n] \) for \( j = 1, 2 \). Let \( V^n \) be an arbitrarily distributed random sequence. Assume that \( (U^n_j(l_j) : l_j \in L_j), (U^n_2(l_2) : l_2 \in L_2) \) and \( V^n \) are independent of each other. Then, \( \delta(\epsilon) \) exists that tends to zero as \( \epsilon \to 0 \) such that

\[
\lim_{n \to \infty} P\{(U^n_1(l_1), U^n_2(l_2), V^n) \in \mathcal{T}_e^n \}
\]

\[
= 0, \quad (10)
\]

if \( r_1 + r_2 < I(U_1; U_2) + I(U_1, U_2; V) - \delta(\epsilon) \).

The proofs of Lemma 1 and Lemma 2 are provided in Appendix A and Appendix B, respectively.

According to Lemma 2, \( P(\mathcal{E}_6) \) tends to zero as \( n \to \infty \) if

\[
(R_1 - R_1) + (R_2 - R_2) < I(U_1; U_2) + I(U_1, U_2; V) - \delta(\epsilon). \quad (11)
\]

By combining (2), (3) and (11), we have shown that \( P(\mathcal{E}) \) tends to zero as \( n \to \infty \) if

\[
R_H > I(V; Y) + \delta(\epsilon''), \quad (12)
\]

\[
R_1 + R_2 > I(U_1; X_1) + I(U_2; X_2) + 2\delta(\epsilon'')
\]

\[
- I(U_1; U_2) - I(U_1, U_2; V) + \delta(\epsilon)
\]

\[
= I(U_1; X_1) + I(U_2; X_2) - I(U_1; U_2)
\]

\[
+ H(X_1 | X_2, U_1) - H(X_1 | X_2, U_1)
\]

\[
- I(U_1, U_2; V) + \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2) + I(U_2; X_2, U_1)
\]

\[
- I(U_1; U_2) + H(X_1 | X_2, U_1)
\]

\[
- H(X_1 | X_2, U_1, U_2) - I(U_1, U_2; V)
\]

\[
+ \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2) + I(U_2; X_2 | U_1)
\]

\[
+ I(U_2; X_1 | X_2, U_1) - I(U_1, U_2; V)
\]

\[
+ \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2) + I(U_2; X_2, U_1)
\]

\[
- I(U_1; U_2; V) + \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2) - I(U_1, U_2; V) + \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2, U_1) - I(U_1, U_2; V)
\]

\[
+ \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2, U_1) - I(U_1, U_2; V)
\]

\[
+ \delta(\epsilon)
\]

\[
= I(U_1; X_1, X_2, U_1) - I(U_1, U_2; V)
\]

\[
+ \delta(\epsilon)
\]

where (13) follows by identifying \( \delta(\epsilon) = 2\delta(\epsilon'') + \delta(\epsilon) \); (14) follows since \( X_2 \to X_1 \to U_1, U_1 \to X_2 \to U_2 \) and \( U_2 \to X_2 \to X_1 \) form three Markov chains; and...
(15) follows since $V \rightarrow (X_1, X_2) \rightarrow (U_1, U_2)$ form a Markov chain.

Note that for some $R_2$ that is large enough for independently decoding and satisfying $D_2$, the events $\mathcal{E}_2$ and $\mathcal{E}_6$ become

$$
\mathcal{E}_2 = \{(U_1^n(\tilde{l}_1), X_1^n) \notin \mathcal{T}_e^{(n)} \text{ for all } \tilde{l}_1 \in [1: 2^{nR_1}]\},
$$

$$
\mathcal{E}_6 = \{(U_1^n(\tilde{l}_1), U_2^n(\tilde{l}_2), V^n(M_H)) \in \mathcal{T}_e^{(n)} \text{ for some } \tilde{l}_1 \in B_1(M_1), \tilde{l}_2 \neq L_1\}.
$$

By the covering lemma, $P(\mathcal{E}_2)$ tends to zero as $n \rightarrow \infty$ if $\tilde{R}_1 > I(U_1; X_1) + \delta(\epsilon)$. Similarly, for some $R_1$ that is large enough for independently decoding and satisfying $D_1$, $P(\mathcal{E})$ tends to zero as $n \rightarrow \infty$ if

$$
\tilde{R}_1 - R_1 < I(U_1; U_2, V) - \delta(\epsilon).
$$

By combining (17) and (18), we have that $P(\mathcal{E})$ tends to zero as $n \rightarrow \infty$ if

$$
R_1 > I(U_1; X_1) + \delta(\epsilon) - I(U_1; U_2, V) + \delta(\epsilon)
$$

$$
= I(U_1; X_1, U_2, V) - I(U_1; U_2, V) + \delta''(\epsilon),
$$

(19)

where (19) follows since $(U_2, V) \rightarrow X_1 \rightarrow U_1$ form a Markov chain and by identifying $\delta''(\epsilon) = \delta(\epsilon) + \delta(\epsilon)$. Similarly, for some $R_1$ that is large enough for independently decoding and satisfying $D_1$, $P(\mathcal{E})$ tends to zero as $n \rightarrow \infty$ if

$$
R_2 > I(U_2; X_2|U_1, V) + \delta''(\epsilon).
$$

(21)

$$
\widehat{p}(\epsilon)\mathcal{T}_e^{(n)}, (U_1^n(L_1), U_2^n(L_2), V^n(M_H), X_1^n, X_2^n, Y^n) \in \mathcal{T}_e^{(n)},
$$

when there is no “error”, Therefore, by the law of total expectation and the typical average lemma, the asymptotic distortions, averaged over the random codebook and encoding, are upper bounded as

$$
\lim_{n \rightarrow \infty} \sup \mathbb{E}(d_j(X_j^n, \hat{X}_j^n))
$$

$$
\leq \lim_{n \rightarrow \infty} \sup d_j(\epsilon) P(\mathcal{E})
$$

$$
+ (1 + \epsilon) \mathbb{E}(d_j(X_j^n, \hat{X}_j^n)\mid P(\mathcal{E}^c))
$$

$$
\leq D_j, \text{ for } j = 1, 2,
$$

(22)

if the inequalities in (12), (16), (20) and (21) are satisfied.

Now, we use time-sharing scheme only between $R_1$ and $R_2$ to establish the achievability of every rate triplet $(R_1, R_2, R_H)$ that satisfies the inequalities in (12), (16), (20) and (21) for some conditional pmf $p(q)p(u_1|x_1, q)p(u_2|x_2, q)$ and functions $\hat{x}_1(u_1, u_2, v)$ and $\hat{x}_2(u_1, u_2, v)$ such that $\mathbb{E}(d_j(X_j^n, \hat{X}_j^n)) < D_j, j = 1, 2$. Finally, using the continuity of mutual information and taking $\epsilon \rightarrow 0$, we can obtain the following inner bound:

Let $(X_1, X_2, Y)$ be a 3-DMS and $d_1(x_1, \hat{x}_1)$ and $d_2(x_2, \hat{x}_2)$ be two distortion measures. A rate triplet $(R_1, R_2, R_H)$ is achievable with distortion pair $(D_1, D_2)$ for distributed lossy source coding with a helper observing $Y$ if

$$
R_1 > I(X_1; U_1|U_2, V, Q),
$$

(23)

$$
R_2 > I(X_2; U_2|U_1, V, Q),
$$

(24)

$$
R_1 + R_2 > I(X_1, X_2; U_1, U_2|V, Q),
$$

(25)

$$
R_H > I(Y; V),
$$

(26)

for some conditional pmf $p(q|x)$ with $|\mathcal{Q}| \leq |\mathcal{Y}| + 1$, and $p(q)p(u_1|x_1, q)p(u_2|x_2, q)$ with $|\mathcal{U}_j| \leq |\mathcal{X}_j| + 4$, $j = 1, 2$, as well as functions $\hat{x}_1(u_1, u_2, v, q)$ and $\hat{x}_2(u_1, u_2, v, q)$ such that $\mathbb{E}(d_j(X_j, \hat{X}_j)) \leq D_j, j = 1, 2$.

IV. PERFORMANCE EVALUATION

In this section, we analyze the performance improvement provided by a helper through the comparison of theoretical inner bound and simulation of a practical wireless communication system.

Fig. 4 compares the inner bound derived in Section III-B with the Slepian-Wolf (S-W) theorem and the Berger-Tung (B-T) inner bound. It is shown that the derived inner bound coincide with Berger-Tung inner bound when the helper rate is equal to 0, and hence it is reasonable and effective to predict the performance improvement by the comparison of theoretical bounds. Interestingly, for arbitrary distortion requirements, the helper can obviously reduce the rate of sensor links as the helper rate increases. Consequently, a helper can eliminate distortion in sensor data, for the given rate of sensor links.

Here, we investigate the effect of helper by evaluating the performance of a practical wireless communication system. As is depicted in Fig. 5, $X_1^n$ and $X_2^n$ are
two correlated binary sources with crossover probability \( \rho \). The helper generates information \( Y^n \) bit by bit as \( Y = X_1 \oplus X_2 \). For simplicity, \( X_1^n, X_2^n \) and \( Y^n \) are randomly punctured by \( \varphi_1, \varphi_2 \) and \( \varphi_h \) as lossy compression, respectively. Then, the punctured sequences are encoded by using a convolutional code (CC) with \( G = ([3, 2, 3]_8) \) and interleaved by an random interleaver \( \Pi \). Finally, the interleaved sequences are encoded by an accumulator (ACC) [17], so as to utilize the principle of turbo code for decoding. After binary phase shift keying (BPSK) modulation, the encoded sequences are sent to the receiver through additive white Gaussian noise (AWGN) channels. Then, the received signals after demodulation are decoded in the joint decoder, by the decoder of ACC (ACC\(^{-1}\)) and the decoder of CC (CC\(^{-1}\)) for the first step. Subsequently, the log-likelihood ratio LLR of all sequences is jointly updated by an extrinsic information exchanger. When iteration achieves maximum time limit, final estimates are made by hard decision.

After 1000 times of simulation with sequence length \( n = 10000 \), the simulation result is illustrated in Fig. 6. Obviously, the helper can significantly reduce the distortion level. It is also remarkable that the system performance with \( R_H = 0 \) matches the curve of a non-helper system. Besides, there is still a gap between theoretical inner bound and simulation results, even if SNR is large enough. The reason is that the encoding and decoding scheme used in simulation cannot fully exploit the joint typicality as in the 3D distributed compress-bin scheme for theoretical proof. Hence, there is a significant potential in the design of coding scheme for WSNs with a helper.

V. CONCLUSION

We have investigated the impact of a helper on the system performance of WSNs. Initially, the system model of WSNs with a helper is formulated as a problem of multiterminal source coding problem with a helper. Then, we raise a 3D distributed compress-bin scheme and derive a inner bound on rate-distortion region for theoretical analysis. Finally, the effect of a helper is studied in depth by comparing the derived inner bound with Berger-Tung inner bound and through simulation. The potential of a helper on performance improvement is verified by the theoretical and simulation results.

APPENDIX A

PROOF OF LEMMA 1

Consider

\[
P\{(\hat{U}_1, \hat{U}_2, \hat{v}) \in \mathcal{T}_e(n)(U_1, U_2, V)\} = \sum_{\hat{v} \in \mathcal{T}_e(n)(V|U_1, U_2)} \sum_{(u_1, u_2) \in \mathcal{T}_e(n)(U_1, U_2)} p(\hat{v}^n) \
\]

\[
\leq P\{(u_1, u_2) \in \mathcal{T}_e(n)(U_1, U_2)\} \cdot |\mathcal{T}_e(n)(V|U_1, U_2)| \
\times 2^{-n(H(V) - \epsilon H(V))} \
\leq P\{(u_1, u_2) \in \mathcal{T}_e(n)(U_1, U_2)\} \
\times 2^{-n(H(V|U_1, U_2) + \epsilon H(V|U_1, U_2))} \
\times 2^{-n(H(V) - \epsilon H(V))} \
\leq 2^{-n(H(V|U_1, U_2) + \epsilon H(V|U_1, U_2))} \
\times 2^{-n(H(V) - \epsilon H(V))} \
\leq 2^{-n(H(V|U_1, U_2) + \epsilon H(V|U_1, U_2) - \delta)} \
\] (27)

where (27) follows by the joint typicality lemma.

APPENDIX B

PROOF OF LEMMA 2

Define the events

\[
\tilde{E}_{l_1, l_2} = \{(U_1^n(l_1), U_2^n(l_2), \tilde{V}^n) \in \mathcal{T}_e(n)\} \
\text{for } (l_1, l_2) \in \mathcal{L}_1 \times \mathcal{L}_2. 
\] (29)
By the union of events bound, the probability of the event of interest can be bounded as
\[
P \left( \bigcup_{(l_1, l_2) \in \mathcal{L}_1 \times \mathcal{L}_2} \tilde{E}_{l_1, l_2} \right) \\
\leq \sum_{(l_1, l_2) \in \mathcal{L}_1 \times \mathcal{L}_2} P(\tilde{E}_{l_1, l_2}).
\] (30)

By Lemma 1,
\[
P(\tilde{E}_{l_1, l_2}) = P\{ (U^n_1(l_1), U^n_2(l_2), \tilde{V}^n) \\
\in T^{(n)}(U_1, U_2, V) \} \\
\leq 2^{-n(I(U_1; U_2) + I(U_1; U_2; V) - \delta(\epsilon))}.\]
(31)

Hence,
\[
\sum_{(l_1, l_2) \in \mathcal{L}_1 \times \mathcal{L}_2} P(\tilde{E}_{l_1, l_2}) \\
\leq 2^{nr_1} . 2^{nr_2} \\
\cdot 2^{-n(I(U_1; U_2) + I(U_1; U_2; V) - \delta(\epsilon))},\] (32)

which tends to zero as \( n \to \infty \) if \( r_1 + r_2 < I(U_1; U_2) + I(U_1; U_2; V) - \delta(\epsilon) \).

ACKNOWLEDGMENTS

This work is funded in part by China Scholarship Council (CSC) and in part by JAIST Core-to-Core Program.

REFERENCES