



HAL
open science

Sparse Random Linear Network Coding for Data Compression in WSNs

Wenjie Li, Francesca Bassi, Michel Kieffer

► **To cite this version:**

Wenjie Li, Francesca Bassi, Michel Kieffer. Sparse Random Linear Network Coding for Data Compression in WSNs. 2016 IEEE International Symposium on Information Theory (ISIT), Jul 2016, Barcelona, Spain. 10.1109/isit.2016.7541795 . hal-01327498

HAL Id: hal-01327498

<https://hal.science/hal-01327498>

Submitted on 6 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sparse Random Linear Network Coding for Data Compression in WSNs

Wenjie Li

Laboratoire des Signaux et Systèmes
CNRS–CentraleSupélec–Univ Paris-Sud
91192, Gif-sur-Yvette, France
Email: wenjie.li@l2s.supelec.fr

Francesca Bassi

ESME-Sudria
94200, Ivry-sur-Seine, France
Email: francesca.bassi@l2s.supelec.fr

Michel Kieffer

Laboratoire des Signaux et Systèmes
CNRS–CentraleSupélec–Univ Paris-Sud
91192, Gif-sur-Yvette, France
Email: michel.kieffer@l2s.supelec.fr

Abstract—This paper addresses the information theoretical analysis of data compression achieved by random linear network coding in wireless sensor networks. A sparse network coding matrix is considered with columns having possibly different sparsity factors. For stationary and ergodic sources, necessary and sufficient conditions are provided on the number of required measurements to achieve asymptotically vanishing reconstruction error. To ensure the asymptotically optimal compression ratio, the sparsity factor can be arbitrary close to zero in absence of additive noise. In presence of noise, a sufficient condition on the sparsity of the coding matrix is also proposed.¹

I. INTRODUCTION

Wireless sensor networks (WSN) [1] are composed by autonomous nodes, with sensing capability of some physical phenomenon (*e.g.* temperature, or pressure). In order to ensure ease of deployment and robustness, the communication between the nodes might need to be performed in absence of designated access points and of a hierarchical structure. At the network layer, dissemination of the measurements to all the nodes can be achieved using an asynchronous protocol based on network coding (NC) [2] in order to reduce the traffic. Each node in the network broadcasts a packet evaluated as the linear combination of its local measurements, and of the packets received from neighboring nodes. Once the sink node has collected enough linearly independent combinations of the network measurements, it can perform decoding by solving a system of linear equations.

In many practical situations, the measurements obtained by the WSN are spatially and temporally correlated. Thus, joint source-network coding can be performed using techniques borrowed from compressive sensing (CS) [3]. The main difficulty comes from the fact that usually, sensor readings are real-valued or quantized and NC operations are done in finite fields. As a consequence, the NC matrix, which plays the role of the sensing matrix in CS, has elements in a finite field and it is difficult to exploit the efficient reconstruction algorithms of CS. A first solution to this problem is proposed in [4], where NC is performed in the real field. This framework is also applied in [5], where data reconstruction is performed progressively to reduce the decoding delay. In [6], linear

combination resulting from real-field NC are quantized before transmission. The price to be paid by all these technique is larger headers and an incompatibility with classical NC.

This paper considers random linear network coding (RLNC) [7], where both quantized measurements and NC coefficients are chosen in the same finite field. With RLNC, the coding coefficients are randomly chosen and sent in each packet header [8]. Usually, NC vectors are sparse: packets reaching the sink are seldom combinations of all packets and depend on the network topology. This allows compressing efficiently the NC vectors [9] and reduces the decoding complexity [10]. Even if the local coding coefficients are independently chosen by each node with the same distribution, the statistical properties and the sparsity of the NC coding vectors reaching the sink may differ significantly.

This paper thus investigates the compression efficiency of RLNC when the NC matrix is assumed to be sparse with rows with possibly different sparsity. The problem can be formulated as the estimation of a source vector θ^N (of N symbols), obtained by acquiring M non-adaptive linear measurements, with $M < N$. In a Bayesian setting, θ is understood as a realization of a random source Θ , whose *a priori* distribution induces compressibility (sparsity and/or correlation between the symbols). The reconstruction problem is formulated in this case as an estimation problem, solvable using standard Bayesian techniques, *e.g.*, Maximum *A Posteriori* (MAP) estimation.

A similar problem has been addressed in [11], where specific correlation patterns (pairwise, cluster) between the elements of Θ are considered and the statistical property of the coding matrix depends on the clusters. Error exponents under MAP decoding are derived in the noiseless case. This problem of compressed sensing over finite field framework has been previously considered in [12], [13]. The vector of measurements at the nodes is interpreted as the compressible source Θ , the network coding matrix as the sensing matrix. In [12], the sensing matrix is considered as uniformly distributed and the prior distribution induces sparsity on θ , whose elements are assumed statistically independent. The error exponent with respect to exact reconstruction is derived with ℓ_0 -norm minimization decoding for noiseless measurements, and with minimum-empirical entropy decoding for noisy measurements.

¹This work has been partly supported by the NEWCOM# NoE and by the iCODE Institute, research project of the IDEX Paris-Saclay.

The work in [13] assumes a known sparsity level of θ and does not consider additive noise. Ideal decoding via ℓ_0 -norm minimization is assumed, and necessary and sufficient conditions for exact recovery are derived as functions of the size of θ , its sparsity level, the number of measurements, and the sparsity of the sensing matrix. Numerical results show that the necessary and sufficient conditions coincide, as the size of θ asymptotically increases.

The major contribution of this work is that we consider a sparse coding matrix with entries of different sparsity level, which is a more realistic hypothesis. No analysis has been previously performed for such model. We assume a Bayesian setting with very mild assumptions on the distribution of the source (Θ is an ergodic process), and we consider MAP decoding of noisy measurements. We aim to derive sufficient conditions on the ratio M/N for almost sure exact recovery of θ , as its size asymptotically increases, and show their convergence to necessary conditions. Our results generalize those presented in [11], [12], and [13]. In addition, we can formally prove the asymptotic convergence of necessary and sufficient conditions on M/N both in the noiseless and noisy settings. Sufficient conditions on the sparsity of the coding matrix in the noisy setting are also provided. These conditions, which depend on the noise distribution, have never been discussed in the previous work, to the best of our knowledge. If no noise is present, the coding matrix could be arbitrarily small to ensure the optimal asymptotical compression ratio.

II. MOTIVATING EXAMPLE

Consider a WSN which topology is shown in Figure 1(a). $S = 29$ wireless sensor nodes are uniformly distributed over a square of unit area with a sink node located at the center. Assume that each sensor Node i attempts to transmit a binary measurement θ_i to the sink using RLNC over the Galois field \mathbb{F}_2 . Time is slotted and each sensor broadcasts a packet or combination of packets in its time slot, starting from the nodes the farthest from the sink. Each node can only communicate with its neighbor nodes at a distance less than 0.25. When Node i performs a combination of packets, its local measurement θ_i is assumed to be always used, whereas a packet received from one of its neighbors has a probability ν to be involved. After N_R communication rounds, the sink is able to build an NC matrix with $M = N_R S_0$ lines and S columns, where S_0 is its number of one-hop neighbors. The sparsity of each column of the NC matrix is then evaluated. Figure 1(b) represents the proportion of non-zero coefficients of each column parametrized in ν averaged over 10^6 independent Monte-Carlo simulations of the network. Only connected networks are considered in the simulations. The average value of M is $17 < S$. Columns associated to nodes which are far from the sink are sparser than columns associated to nodes close to the sink. Moreover, a small value of ν makes the overall coding matrix sparser. This justifies our study considering coding matrices with sparsity varying among columns.

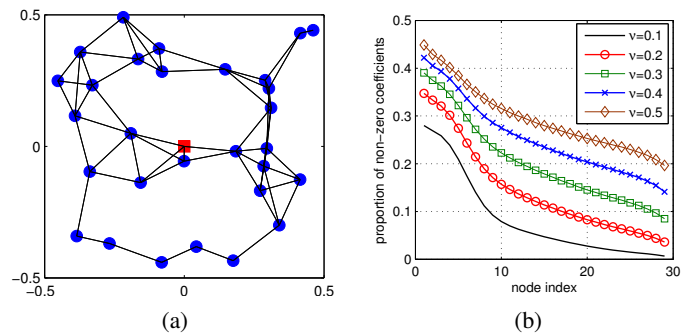


Figure 1. (a) example network topology ; (b) evolution of the proportion of non-zero NC coefficient as a function of the node index sorted by increasing distance to the sink, for $\nu \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$.

III. SYSTEM MODEL AND PROBLEM SETUP

In what follows, sans-serif is used to denote random quantities while deterministic quantities are with serif fonts. Matrices are in bold-face upper-case letters. A length n vector is in bold-face lower-case with a superscript n . Calligraphic font denotes set, except \mathcal{H} , which denotes the entropy rate. All logarithms are in base 2.

A. System Model

Consider a WSN that consists of S sensor nodes and each sensor node has T measurements, denote $N = ST$. The measurements are quantized before transmission and represented by a length- N random source vector Θ^N with entries taking values in the Galois Field \mathbb{F}_Q of size Q . Let $\theta^N \in \mathbb{F}_Q^N$ be a realization of Θ^N . The probability mass function (pmf) associated with Θ^N is denoted by $p_{\Theta^N}(\theta^N)$, or $p(\theta^N)$, when there is no ambiguity. The correlated source Θ^N is assumed to be stationary and ergodic with an entropy rate $\mathcal{H}(\Theta)$.

Considering the random coding matrix $\mathbf{A} \in \mathbb{F}_Q^{M \times N}$, assumed to be independent of Θ^N , the sink gets the vector of measurements $\mathbf{y}^M = \mathbf{A}\theta^N + \mathbf{u}^M$, where $\mathbf{u}^M \in \mathbb{F}_Q^M$ can represent some additive noise due to some computation error or Byzantine attacks, etc. Each entry of \mathbf{u}^M is assumed to be iid with pmf p_u . Denote \mathbf{A}_i as the i -th column of \mathbf{A} . The entries of \mathbf{A}_i are assumed to be iid in \mathbb{F}_Q with pmf

$$p_{\mathbf{A}_i}(0) = 1 - \gamma_i, \quad p_{\mathbf{A}_i}(q) = \frac{\gamma_i}{Q-1} \text{ for } q \in \mathbb{F}_Q \setminus \{0\} \quad (1)$$

where γ_i is the sparsity factor of \mathbf{A}_i . The sparsity factors γ_i and γ_j of two different columns \mathbf{A}_i and \mathbf{A}_j may be different. Define $\gamma = \min_{i=1, \dots, N} \gamma_i$ as the minimum sparsity factor.

B. MAP Decoding

With perfect knowledge of the realizations \mathbf{y}^M and \mathbf{A} , the decoder can reconstruct θ^N using an MAP estimation approach. The estimate $\hat{\theta}^N$ is evaluated as $\hat{\theta}^N = \arg \max_{\theta^N \in \mathbb{F}_Q^N} p(\theta^N | \mathbf{y}^M, \mathbf{A})$, where

$$\begin{aligned} p(\theta^N | \mathbf{y}^M, \mathbf{A}) &\propto p(\theta^N, \mathbf{y}^M, \mathbf{A}) \\ &= \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p(\theta^N) p(\mathbf{u}^M) p(\mathbf{A}) p(\mathbf{y}^M | \theta^N, \mathbf{u}^M, \mathbf{A}). \end{aligned} \quad (2)$$

Notice that $p(\mathbf{y}^M | \boldsymbol{\theta}^N, \mathbf{u}^M, \mathbf{A}) = \mathbb{1}_{\mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M}$ is an indicator function, for fixed \mathbf{y}^M , \mathbf{A} , and $\boldsymbol{\theta}^N$, there is exactly one vector \mathbf{u}^M such that $\mathbf{u}^M = \mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N$, thus

$$p(\boldsymbol{\theta}^N | \mathbf{y}^M, \mathbf{A}) \propto p(\boldsymbol{\theta}^N) p_{\mathbf{u}^M}(\mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N) p(\mathbf{A}). \quad (3)$$

A decoding error happens if there exists a vector $\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}$ such that

$$p_{\boldsymbol{\theta}^N}(\boldsymbol{\varphi}^N) p_{\mathbf{u}^M}(\mathbf{y}^M - \mathbf{A}\boldsymbol{\varphi}^N) \geq p_{\boldsymbol{\theta}^N}(\boldsymbol{\theta}^N) p_{\mathbf{u}^M}(\mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N).$$

An alternative way to state the error event is: For a given realization $\boldsymbol{\theta}^N$, which implies the realization $\mathbf{u}^M = \mathbf{y}^M - \mathbf{A}\boldsymbol{\theta}^N$, there exists a pair $(\boldsymbol{\varphi}^N, \mathbf{v}^M) \in \mathbb{F}_Q^N \times \mathbb{F}_Q^M$ such that

$$\begin{cases} \boldsymbol{\varphi}^N \neq \boldsymbol{\theta}^N, \\ \mathbf{A}\boldsymbol{\varphi}^N + \mathbf{v}^M = \mathbf{y}^M = \mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M, \\ p(\boldsymbol{\varphi}^N) p(\mathbf{v}^M) \geq p(\boldsymbol{\theta}^N) p(\mathbf{u}^M). \end{cases} \quad (4)$$

Let $\Pr\{e | \boldsymbol{\theta}^N, \mathbf{u}^M\}$ denote the conditional error probability for fixed $\boldsymbol{\theta}^N$ and \mathbf{u}^M . The average error probability P_e is

$$P_e = \sum_{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N} \sum_{\mathbf{u}^M \in \mathbb{F}_Q^M} p(\boldsymbol{\theta}^N, \mathbf{u}^M) \Pr\{e | \boldsymbol{\theta}^N, \mathbf{u}^M\}. \quad (5)$$

Our objective is to derive necessary and sufficient conditions for asymptotically vanishing P_e .

C. Necessary Condition

Proposition 1 (Necessary condition). *Consider some arbitrary small $\delta \in \mathbb{R}^+$. For $N \rightarrow \infty$, the necessary conditions for $P_e < \delta$ are*

$$H(p_u) < \log Q, \quad (6)$$

$$\frac{M}{N} > \frac{\mathcal{H}(\Theta) - \varepsilon}{\log Q - H(p_u)}, \quad (7)$$

where $\varepsilon \in \mathbb{R}^+$ is an arbitrary small constant.

Proposition 1 extends results obtained in [13] to the noisy case. The proof involves a simple application of Fano's inequality, see [14] for the detail.

D. Sufficient Condition

Proposition 2 (Sufficient condition). *Consider a coding matrix with minimum sparsity factor γ . For any $\delta > 0$, there exist small positive real numbers ε , ξ , and integers N_δ , M_ε such that $\forall N > N_\delta$ and $\forall M > M_\varepsilon$, if the following conditions hold:*

$$H(p_u) < \log Q - \xi, \quad (8)$$

$$\gamma > 1 - 2^{-H(p_u) - \varepsilon}, \quad (9)$$

$$\frac{M}{N} > \frac{\mathcal{H}(\Theta) + \varepsilon}{\log Q - H(p_u) - \xi}, \quad (10)$$

then one has $P_e \leq \delta$ using MAP decoding. As $N \rightarrow \infty$ and $M \rightarrow \infty$, ε and ξ can be arbitrary close to zero.

The proof detail of Proposition 2 is given in Section IV.

In the noiseless case, the necessary condition in Proposition 1 and the sufficient condition in Proposition 2 asymptotically coincide, since γ can be chosen arbitrarily small. This

confirms the numerical results obtained in [13]. In the noisy case, the difference between the two conditions comes from the constraint linking γ and the entropy of the communication noise. In the analysis of the necessary conditions, the structure of \mathbf{A} was not considered and no condition on γ has been obtained. The lower bound on γ implies that \mathbf{A} should be dense enough to fight against the noise.

IV. PROOF OF PROPOSITION 2

A. Upper Bound of the Error Probability

Proposition 3. *Under MAP decoding, the asymptotic ($N \rightarrow \infty$) probability of error can be upper bounded as*

$$P_e \leq 2^{-NE_1(\alpha)} + 2^{-NE_2(\alpha)} + 2\varepsilon, \quad (11)$$

where $\varepsilon \in \mathbb{R}^+$ is an arbitrary small constant. $E_1(\alpha)$ and $E_2(\alpha)$ are defined as

$$E_1(\alpha) = -\frac{M}{N} (H(p_u) + \log(1 - \gamma) + \varepsilon) - H_2(\alpha) - \alpha \log(Q - 1) - \frac{\log(\alpha N)}{N}, \quad (12)$$

$$E_2(\alpha) = -\frac{M}{N} \log \left(Q^{-1} + \left(1 - \frac{\gamma}{1 - Q^{-1}} \right)^{\lceil \alpha N \rceil} (1 - Q^{-1}) \right) - \mathcal{H}(\Theta) - \frac{M}{N} (H(p_u) + \varepsilon) - \varepsilon, \quad (13)$$

with $\alpha \in \mathbb{R}^+$ and $\alpha < 0.5$.

Proof. Weak typicality is instrumental in the following proofs. For any positive real number ε and some integer $N > 0$, the weakly typical set $\mathcal{A}_{[\Theta]^\varepsilon}^N \subset \mathbb{F}_Q^N$ for a stationary and ergodic source Θ^N is $\mathcal{A}_{[\Theta]^\varepsilon}^N = \{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N : \left| -\frac{1}{N} \log p(\boldsymbol{\theta}^N) - \mathcal{H}(\Theta) \right| \leq \varepsilon\}$. Define $\mathcal{A}_{[u]^\varepsilon}^M$ for the noise vector \mathbf{u}^M similarly. Recall that the entries of \mathbf{u}^M are uncorrelated, so $\mathcal{H}(\mathbf{u}) = H(p_u)$. Thanks to Shannon-McMillan-Breiman theorem [15, Sec. 16.8], the pmf of the general stationary and ergodic source converges. In other words, for any $\varepsilon > 0$, there exists N_ε and M_ε such that for all $N > N_\varepsilon$ and $M > M_\varepsilon$, one has $\Pr\{\boldsymbol{\theta}^N \in \mathcal{A}_{[\Theta]^\varepsilon}^N\} \geq 1 - \varepsilon$ and $\Pr\{\mathbf{u}^M \in \mathcal{A}_{[u]^\varepsilon}^M\} \geq 1 - \varepsilon$. We can make ε arbitrary close to zero as $N \rightarrow \infty$ and $M \rightarrow \infty$. Define $\mathcal{U} = \mathcal{A}_{[\Theta]^\varepsilon}^N \times \mathcal{A}_{[u]^\varepsilon}^M$ and \mathcal{U}^c such that $\mathcal{U} \cup \mathcal{U}^c = \mathbb{F}_Q^N \times \mathbb{F}_Q^M$. One has

$$\begin{aligned} P_e &= \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \cup \mathcal{U}^c} p(\boldsymbol{\theta}^N) p(\mathbf{u}^M) \Pr\{e | \boldsymbol{\theta}^N, \mathbf{u}^M\} \\ &\leq \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} p(\boldsymbol{\theta}^N) p(\mathbf{u}^M) \Pr\{e | \boldsymbol{\theta}^N, \mathbf{u}^M\} + 2\varepsilon, \end{aligned} \quad (14)$$

which comes from the fact that $\Pr\{e | \boldsymbol{\theta}^N, \mathbf{u}^M\} \leq 1$ and $\Pr\{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}^c\} \leq 2\varepsilon$.

Since \mathbf{A} is generated randomly, define the random event $\mathcal{E}(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M) = \{\mathbf{A}\boldsymbol{\theta}^N + \mathbf{u}^M = \mathbf{A}\boldsymbol{\varphi}^N + \mathbf{v}^M\}$, where $(\boldsymbol{\theta}^N, \mathbf{u}^M)$ is the realization of the environment state, and $(\boldsymbol{\varphi}^N, \mathbf{v}^M)$ is the potential reconstruction result. Conditioned

on $(\boldsymbol{\theta}^N, \mathbf{u}^M)$, $\Pr\{e \mid \boldsymbol{\theta}^N, \mathbf{u}^M\}$ is the probability of the union of the events $\mathcal{E}(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M)$ with all the parameter pairs $(\boldsymbol{\varphi}^N, \mathbf{v}^M) \in \mathbb{F}_Q^N \times \mathbb{F}_Q^M$ such that $\boldsymbol{\varphi}^N \neq \boldsymbol{\theta}^N$ and $p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M) \geq p(\boldsymbol{\theta}^N)p(\mathbf{u}^M)$, see (4). One has

$$\begin{aligned} & \Pr\{e \mid \boldsymbol{\theta}^N, \mathbf{u}^M\} \\ &= \Pr\left\{ \bigcup_{\substack{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}, \\ \mathbf{v}^M \in \mathbb{F}_Q^M: p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M) \geq p(\boldsymbol{\theta}^N)p(\mathbf{u}^M)}} \mathcal{E} \right\} \\ &\stackrel{(a)}{\leq} \sum_{\substack{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}, \\ \mathbf{v}^M \in \mathbb{F}_Q^M}} \Phi(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M) \Pr\{\mathcal{E}\} \\ &\stackrel{(b)}{\leq} \sum_{\substack{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}, \\ \mathbf{v}^M \in \mathbb{F}_Q^M}} \left(\frac{p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M)}{p(\boldsymbol{\theta}^N)p(\mathbf{u}^M)} \right)^s \Pr\{\mathcal{E}\} \quad (15) \end{aligned}$$

where (a) is by applying the union bound. An indicator function is also introduced in (a), $\Phi(\boldsymbol{\theta}^N, \mathbf{u}^M; \boldsymbol{\varphi}^N, \mathbf{v}^M) = 1$ if $p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M) \geq p(\boldsymbol{\theta}^N)p(\mathbf{u}^M)$, otherwise it is equal to 0. The idea of (b) comes from a part of Gallager's derivation of error exponents in [16, Sec. 5.6], the parameter $s \in \mathbb{R}$ and $0 < s \leq 1$. Introducing (15) in (14) and setting $s = 1$ yields

$$P_e \leq \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N \setminus \{\boldsymbol{\theta}^N\}} \sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M) \Pr\{\mathcal{E}\} + 2\varepsilon. \quad (16)$$

In (16), $\Pr\{\mathcal{E}\} = \Pr\{\mathbf{A}\boldsymbol{\mu}^N = \mathbf{s}^M\}$ with $\boldsymbol{\mu}^N = \boldsymbol{\varphi}^N - \boldsymbol{\theta}^N \in \mathbb{F}_Q^N \setminus \{\mathbf{0}\}$, and $\mathbf{s}^M = \mathbf{u}^M - \mathbf{v}^M \in \mathbb{F}_Q^M$. This probability depends on $d_1 = \|\boldsymbol{\mu}^N\|_0$ and $d_2 = \|\mathbf{s}^M\|_0$. Both d_1 and d_2 are integers such that $1 \leq d_1 \leq N$ and $0 \leq d_2 \leq M$. Define

$$f(d_1, d_2) = \Pr\{\mathbf{A}\boldsymbol{\mu}^N = \mathbf{s}^M \mid \|\boldsymbol{\mu}^N\|_0 = d_1, \|\mathbf{s}^M\|_0 = d_2\}. \quad (17)$$

Lemma 1. *The function $f(d_1, d_2)$ is non-increasing in d_2 for a given d_1 and*

$$f(d_1, d_2) \leq f(d_1, 0) = \left(Q^{-1} + \left(1 - \frac{\gamma}{1 - Q^{-1}}\right)^{d_1} (1 - Q^{-1}) \right)^M. \quad (18)$$

Moreover $f(d_1, 0)$ is non-increasing in d_1 and

$$f(d_1, 0) \leq f(1, 0) = (1 - \gamma)^M. \quad (19)$$

See [14] for the proof details. Using Lemma 1, (16) can be expressed as

$$\begin{aligned} P_e &\stackrel{(a)}{\leq} \sum_{d_1=1}^N \sum_{d_2=0}^M \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1 \\ \mathbf{v}^M \in \mathbb{F}_Q^M: \|\mathbf{u}^M - \mathbf{v}^M\|_0 = d_2}} p(\boldsymbol{\varphi}^N)p(\mathbf{v}^M) f(d_1, d_2) + 2\varepsilon \\ &\stackrel{(b)}{\leq} \sum_{d_1=1}^N \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p(\boldsymbol{\varphi}^N) f(d_1, 0) \sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p(\mathbf{v}^M) + 2\varepsilon \quad (20) \end{aligned}$$

where (a) is by the classification of $\boldsymbol{\varphi}^N$ and \mathbf{v}^M according to the ℓ_0 norm of their difference with $\boldsymbol{\theta}^N$ and \mathbf{u}^M respectively and (b) is obtained using the bound (18) and using $\sum_{\mathbf{v}^M \in \mathbb{F}_Q^M} p(\mathbf{v}^M) = 1$. From (20), one performs a splitting to bound $f(d_1, 0)$ in different cases,

$$\begin{aligned} P_e &\leq \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1} p(\boldsymbol{\varphi}^N) f(1, 0) \\ &\quad + \sum_{d_1=\lceil \alpha N \rceil}^N \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p(\boldsymbol{\varphi}^N) f(\lceil \alpha N \rceil, 0) + 2\varepsilon, \quad (21) \end{aligned}$$

The parameter α is a positive real number with $0 < \alpha < 0.5$. The way to choose α will be discussed in Section IV-B. For the first term $P_{\mathcal{U}_1}(\alpha)$ in (21),

$$\begin{aligned} P_{\mathcal{U}_1}(\alpha) &= f(1, 0) \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p(\boldsymbol{\varphi}^N) \\ &\stackrel{(a)}{=} (1 - \gamma)^M \sum_{\mathbf{u}^M \in \mathcal{A}_{[\mathbf{u}]_\varepsilon}^M} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N} p(\boldsymbol{\varphi}^N) \sum_{\substack{\boldsymbol{\theta}^N \in \mathcal{A}_{[\boldsymbol{\theta}]_\varepsilon}^N \\ \|\boldsymbol{\theta}^N - \boldsymbol{\varphi}^N\|_0 = d_1}} 1 \\ &\stackrel{(b)}{\leq} (1 - \gamma)^M \sum_{\mathbf{u}^M \in \mathcal{A}_{[\mathbf{u}]_\varepsilon}^M} \sum_{d_1=1}^{\lfloor \alpha N \rfloor} 2^{NH_2(\frac{d_1}{N})} (Q - 1)^{d_1} \\ &\stackrel{(c)}{\leq} (1 - \gamma)^M \cdot \left| \mathcal{A}_{[\mathbf{u}]_\varepsilon}^M \right| \cdot \alpha N \cdot 2^{NH_2(\alpha)} (Q - 1)^{\alpha N} \\ &\stackrel{(d)}{\leq} 2^{-NE_1(\alpha)}, \quad (22) \end{aligned}$$

where (a) is by changing the order of summation; (b) is by ignoring $\boldsymbol{\theta}^N \in \mathcal{A}_{[\boldsymbol{\theta}]_\varepsilon}^N$ and

$$\sum_{\substack{\boldsymbol{\theta}^N \in \mathbb{F}_Q^N: \\ \|\boldsymbol{\theta}^N - \boldsymbol{\varphi}^N\|_0 = d_1}} 1 = \binom{N}{d_1} (Q - 1)^{d_1} \leq 2^{NH_2(\frac{d_1}{N})} (Q - 1)^{d_1},$$

where $H_2(p)$ denotes the entropy of a Bernoulli- p source; (c) is due to the fact that $H_2(\frac{d_1}{N})$ is increasing in d_1 as $d_1 \leq \lfloor \alpha N \rfloor < N/2$; (d) comes from [15, Theorem 3.1.2], the upper bound of the size of the set, $\left| \mathcal{A}_{[\mathbf{u}]_\varepsilon}^M \right| \leq 2^{M(H(p_u) + \varepsilon)}$ for $M > M_\varepsilon$. One gets finally $E_1(\alpha)$, as defined in (12). Now we turn to $P_{\mathcal{U}_2}(\alpha)$,

$$\begin{aligned} P_{\mathcal{U}_2}(\alpha) &= \sum_{d_1=\lceil \alpha N \rceil}^N \sum_{\substack{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U} \\ \boldsymbol{\varphi}^N \in \mathbb{F}_Q^N: \|\boldsymbol{\varphi}^N - \boldsymbol{\theta}^N\|_0 = d_1}} p(\boldsymbol{\varphi}^N) f(\lceil \alpha N \rceil, 0) \\ &\stackrel{(a)}{\leq} \sum_{(\boldsymbol{\theta}^N, \mathbf{u}^M) \in \mathcal{U}} \sum_{\boldsymbol{\varphi}^N \in \mathbb{F}_Q^N} p(\boldsymbol{\varphi}^N) f(\lceil \alpha N \rceil, 0) \\ &= \left| \mathcal{A}_{[\boldsymbol{\theta}]_\varepsilon}^N \right| \left| \mathcal{A}_{[\mathbf{u}]_\varepsilon}^M \right| f(\lceil \alpha N \rceil, 0) \stackrel{(b)}{\leq} 2^{-NE_2(\alpha)}, \quad (23) \end{aligned}$$

where (a) is by ignoring $\|\varphi^N - \theta^N\|_0 = d_1$, (b) is by the upper bounds of $\left| \mathcal{A}_{[\Theta]^\varepsilon}^N \right|$ and $\left| \mathcal{A}_{[u]^\varepsilon}^M \right|$, and $E_2(\alpha)$ is defined in (13). Combining (21-23) completes the proof. \square

B. Proof of the Sufficient Condition

Both $2^{-NE_1(\alpha)}$ in (22) and $2^{-NE_2(\alpha)}$ in (23) need to be vanishing for increasing N and M . The exponent of each term is considered separately. If $E_1(\alpha) > 0$, for any $\tau_1 \in \mathbb{R}^+$ arbitrarily small, $\exists N_{\tau_1}$ such that $\forall N > N_{\tau_1}$, one has $2^{-NE_1(\alpha)} < \tau_1$. If $H(p_u) + \log(1-\gamma) + \varepsilon \geq 0$, $E_1(\alpha)$ is negative, thus one should first have $H(p_u) + \log(1-\gamma) + \varepsilon < 0$, leading to (9). With this condition, $E_1(\alpha) > 0$ leads to

$$\frac{M}{N} > \frac{H_2(\alpha) + \alpha \log(Q-1) + \frac{\log(\alpha N)}{N}}{\log \frac{1}{1-\gamma} - H(p_u) - \varepsilon}. \quad (24)$$

Similarly, if $E_2(\alpha) > 0$, for any $\tau_2 \in \mathbb{R}^+$ arbitrarily small, $\exists N_{\tau_2} \in \mathbb{N}^+$ such that $\forall N > N_{\tau_2}$, one has $2^{-NE_2(\alpha)} < \tau_2$. Since $0 < \gamma \leq 1 - Q^{-1}$, one gets $0 \leq 1 - \frac{\gamma}{1-Q^{-1}} < 1$ and $\lim_{N \rightarrow \infty} \left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N \rceil} = 0$. Thus for $\sigma \in \mathbb{R}^+$ arbitrarily small, $\exists N_\sigma$ such that $\forall N > N_\sigma$,

$$\left(1 - \frac{\gamma}{1-Q^{-1}}\right)^{\lceil \alpha N \rceil} (1 - Q^{-1}) < \sigma Q^{-1}. \quad (25)$$

Hence $E_2(\alpha)$ in (13) can be lower bounded by

$$E_2(\alpha) > -\mathcal{H}(\Theta) - \frac{M}{N} (H(p_u) - \log Q + \xi) - \varepsilon, \quad (26)$$

for $N > N_\sigma$, with

$$\xi = \log(1 + \sigma) + \varepsilon. \quad (27)$$

If this lower bound is positive, then $E_2(\alpha)$ is positive. Again, if $H(p_u) - \log Q + \xi \geq 0$, one obtains a negative lower bound for $E_2(\alpha)$ from (26). Thus, one deduces (8) in Proposition 2. From (8), to get a positive lower bound for (26), one should have (10) in Proposition 2, with $\varepsilon \rightarrow 0$ and $\xi \rightarrow 0$ as $N \rightarrow \infty$ and $M \rightarrow \infty$. The compression ratio M/N is lower bounded by both (10) and (24). The value of α should be chosen such that the required lower bound on M/N is minimum. One may compare the necessary condition (7) with (10) and (24) respectively. Obviously, (10) is similar to (7), since both ξ and ε can be made arbitrarily close to 0 as $N \rightarrow \infty$. A good value for α has thus to be such that

$$\frac{H_2(\alpha) + \alpha \log(Q-1) + \frac{\log(\alpha N)}{N}}{\log \frac{1}{1-\gamma} - H(p_u) - \varepsilon} \leq \frac{\mathcal{H}(\Theta) + \varepsilon}{\log Q - H(p_u) - \xi}. \quad (28)$$

The function $H_2(\alpha) + \alpha \log(Q-1)$ is increasing when $\alpha \in]0, 0.5[$ and tends to 0 as $\alpha \rightarrow 0$. The term $\log(\alpha N)/N$ is also negligible for N large. Thus, there always exists some α satisfying (28). Since the speed of convergence of ξ is affected by α , we choose the largest α that satisfies (28). Then the sufficient condition (10) for M/N is obtained.

From (11), one may conclude that $P_e \leq \tau_1 + \tau_2 + 2\varepsilon$. To ensure $P_e < \delta$, we should choose τ_1 , τ_2 , and ε to satisfy $\tau_1 + \tau_2 + 2\varepsilon < \delta$. Then a proper value of σ , which depends on

τ_2 and ε , can be chosen. At last, ξ is obtained from (27). With these well determined parameters, if all the three conditions in Proposition 2 hold, there exists integers N_ε , N_{τ_1} , N_{τ_2} , and N_σ , such that for any $N > N_\delta = \max\{N_\varepsilon, N_{\tau_1}, N_{\tau_2}, N_\sigma\}$ one has $P_e < \delta$.

V. CONCLUSIONS

This paper considers RLNC for data compression in a WSN. An asymptotically sufficient condition on the compression ratio for reliable recovery at some sink of the node measurement vector is obtained. Since necessary and sufficient conditions asymptotically converge, the MAP decoder achieves the optimum lower bound of the compression ratio, even in the case of sparse coding matrices. Several previous results have been generalized by considering a stationary and ergodic source model. The choice of the sparsity factor of the coding matrix depends on the noise.

REFERENCES

- [1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.
- [2] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [3] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [4] S. Feizi and M. Médard, "A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing," in *Proc. 49th Allerton Conference*, 2011, pp. 1048–1054.
- [5] M. Leinonen, M. Codreanu, and M. Juntti, "Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks," *IEEE Trans. Wireless Commu.*, vol. 14, no. 3, pp. 1622–1635, 2015.
- [6] M. Nabae and F. Labeau, "Quantized network coding for correlated sources," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–17, 2014.
- [7] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.
- [8] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. 41-st Allerton Conference*, Oct. 2003.
- [9] M. Jafari, L. Keller, C. Fragouli, and K. Argyraki, "Compressed network coding vectors," in *Proc. IEEE Intl. Symp. on Inf. Theory*, Seoul, Korea, 2009, pp. 109–113.
- [10] F. Bassi, C. Liu, L. Iwaza, and M. Kieffer, "Compressive linear network coding for efficient data collection in wireless sensor networks," in *Proc. 20th EUSIPCO*, Bucharest, Romania, 2012, pp. 714–718.
- [11] K. Rajawat, C. Alfonso, and G. Giannakis, "Network-compressive coding for wireless sensors with correlated data," *IEEE Trans. on Wireless Communications*, vol. 11, no. 12, pp. 4264–4274, 2012.
- [12] S. C. Draper and S. Malekpour, "Compressed sensing over finite fields," in *Proc. Intl. Symp. Inf. Theory*, Seoul, Korea, 2009, pp. 669–673.
- [13] J.-T. Seong and H.-N. Lee, "Necessary and sufficient conditions for recovery of sparse signals over finite fields," *IEEE Communications Letters*, vol. 17, no. 10, pp. 1976–1979, 2013.
- [14] W. Li, F. Bassi, and M. Kieffer, "Robust bayesian compressed sensing over finite fields: asymptotic performance analysis," *arXiv preprint arXiv:1401.4313*, 2014.
- [15] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.
- [16] R. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons., 1968.