TRANSMISSION OF CORRELATED INFORMATION SOURCES WITH NETWORK CODING

Hyunggon Park, Nikolaos Thomos‡ and Pascal Frossard‡*

*Multimedia Communications and Networking Laboratory, Ewha Womans University, Seoul, Korea.
‡Signal Processing Laboratory (LTS4), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
hyunggon.park@ewha.ac.kr, {nikolaos.thomos, pascal.frossard}@epfl.ch

ABSTRACT

This paper addresses the problem of the distributed delivery of correlated data sources with help of network coding. Network coding provides an alternative to routing algorithms and offers improved system performance, robustness and throughput, with no need of deploying sophisticated routing strategies. However, the performance is directly driven by the number of innovative data packets that reach the receiver. If the number of received innovative data packets is significantly small, the decoder cannot perfectly recover the transmitted information. However, we show that the correlation between the data sources can be used at decoder for effective approximate decoding. We analytically investigate the impact of the network coding algorithm, and in particular, of the size of finite fields on the decoding performance. Then, we determine an optimal field size that minimizes the expected decoding error, which represents a trade-off between quantization of the source data and probability of decoding error. The network coding with approximate decoding algorithm is implemented in illustrative multimedia streaming and sensor network applications. In both cases, the experimental results confirm the field size analysis and illustrate the effectiveness of approximate decoding of correlated data.

1. INTRODUCTION

The rapid developments of sensor networks has triggered important research efforts that study the design of low complexity sensing strategies and efficient solutions for information delivery. Since it is often difficult to achieve and maintain the coordination among sensors, the transmission of information from the sensors has typically to be performed in a distributed manner on ad-hoc or overlay mesh network topologies. Network coding [1] has been recently proposed as a method to build efficient distributed delivery algorithms in networks with path and source diversity. It is based on the paradigm, where the network nodes are allowed to perform basic processing operations on information streams. The network nodes can combine information packets and transmit the combined data to the next network nodes. When the decoder receives enough data, it recovers the original information by performing inverse operations (e.g., Gaussian elimination for linear combinations). Such a strategy permits to improve the throughput of the system and to approach better max-flow min-cut limit of networks [2, 3]. It enhances the robustness to data loss and reduces the need for coordination in the transmission of data in overlay networks compared to classical routing and scheduling algorithms. In practice, random linear network coding (RLNC) [4], where coding in the network is based on a random selection of coefficients, is often the preferred network coding solution for the distributed delivery of time-sensitive multimedia information [5]. Using RLNC enables distributed delivery, as each node can act independently with no need for central coordination.

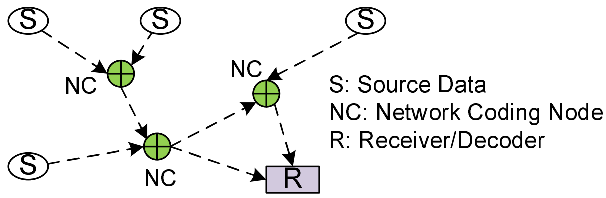

Figure 1: A distributed data transmission system.

We focus on the distributed transmission of correlated data sources with network coding techniques, which is illustrated in Fig. 1. Correlated data can be sources having *external* correlation (e.g., data measured from different locations in sensor networks) or *intrinsic* redundancy (e.g., images in a video sequence). The transmission of correlated sources is generally studied in the framework of distributed coding [6], where sources are encoded by systematic channel encoders and eventually jointly decoded [7, 8]. This choice, however, does not fully exploit the network diversity. Moreover, in the proposed approach, each sensor does not need to know the correlation information, which may enable the proposed solutions to take into account more general scenarios. Thus, network coding is a natural solution to the transmission of correlated data over networks with diversity [9], where it leads to efficient distributed algorithms. However, due to the source and network dynamics, there is no guarantee that each node receives enough useful packets for successful data recovery. This becomes even more critical if applications are delay-sensitive, as delayed packets are discarded due to timing constraints. Thus, it is essential to have a methodology that enables the recovery of the original data with a good accuracy, when the number of innovative packets[1] is not sufficient for perfect decoding.

Since the encoding and decoding processes in each node of RLNC are based on linear operations (e.g., weighted linear combinations, inverse of linear matrix, etc.) in finite algebraic fields, the original data can be approximately recovered

This work was supported in part by the Swiss National Science Foundation grants 200021-118230, PZ00P2-121906, and in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0009717). This work was performed while the first author was with EPFL.

[1] A packet is referred to as *innovative* if it increases the rank of the coding coefficient matrix.

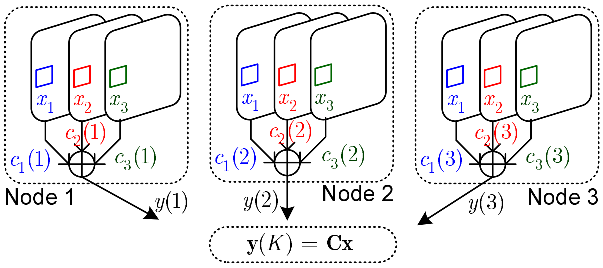

Figure 2: An illustrative example for network coding.

with help of regularization techniques such as Tikhonov regularization [10]. While regularization techniques provide a closed form solution and can be used in a general case, it may result in significantly unreasonable approximations [10]. In this paper, we propose to use the correlation between the sources to design *approximate decoding* algorithms when the number of packets is insufficient. We show that the use of data correlation, such as external correlation or intrinsic redundancy at decoding can lead to an efficient solution for data recovery.

The information about correlation provides additional constraints in the decoding process, such that well-known approaches for matrix inversion (e.g., Gaussian elimination) can be efficiently used. We show analytically that the use of correlation leads to a better data recovery, or equivalently, that the proposed approximate decoding solution results in improved decoding performance. Moreover, we analyze the impact of the accuracy of the correlation information on the decoding performance, since the correlation information is usually obtained from estimation in practice. Our analysis shows that more accurate correlation information leads to better performance in the approximate decoding. We then analyze the influence of the network coding strategy, and in particular, of the choice of the finite field size (i.e., Galois Field (GF) size) on the performance of the approximate decoding. We demonstrate that the GF size should be selected by considering the tradeoff between source approximation and decoding performance. Specifically, the quantization error of the source data decreases with the coding GF size, while the decoding error probability increases with the field size. We show that there is an optimal value for the GF size when approximate decoding is enabled at the receivers. Finally, we illustrate the performance of the network coding algorithm with the approximate decoding on two types of correlated data, i.e., seismic data (external correlation) and video sequences (intrinsic correlation). We demonstrate the results of the finite field size analysis and show that the approximate decoding leads to efficient reconstruction when the correlation information is used during decoding.

This paper is organized as follows. In Section 2, we describe the network coding framework considered in this paper. The influence of the size of the finite field is studied in Section 3. In Section 4, we provide illustrative examples that show how the proposed approach can be implemented in video delivery or sensor networks applications, and conclusions are drawn in Section 5.

2. PROPOSED FRAMEWORK

In this section, we describe the framework considered in this paper and present the encoding and decoding strategies. Let $x_1, \ldots, x_N$ be N non-negative correlated original data, where $x_n \in \mathscr{X}$ for $1 \leq n \leq N$. $\mathscr{X}$ denotes an alphabet size of x_n. In RLNC, a node k transmits packets $y(k) = \sum_{n=1}^{N} c_n(k) x_n$, which is a linear combination of x_n with weights $c_n(k)$ randomly chosen from GF(2^r). These packets are transmitted to other nodes. Hence, the GF size is determined by r. The nodes are distributed over a network (e.g., adhoc network). We assume that $|\mathscr{X}| \leq 2^r$. An illustrative example of the coding process in the case where $N = 3$ is shown in Fig. 2.

If K innovative (i.e., linearly independent) packets, $y(1), \ldots, y(K)$, are available, the following linear system $\mathbf{y}(K) = \mathbf{C}\mathbf{x}$ can be formed[2]:

$$\begin{bmatrix} y(1) \\ \vdots \\ y(K) \end{bmatrix} = \begin{bmatrix} c_1(1) & \ldots & c_N(1) \\ \vdots & \ddots & \vdots \\ c_1(K) & \ldots & c_N(K) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}. \quad (1)$$

The goal of decoding is therefore to estimate $\hat{\mathbf{x}}$ from the received $\mathbf{y}(K)$ in (1). If $K = N$, $\hat{\mathbf{x}}$ can be uniquely determined as $\mathbf{x}$ based on well-known approaches such as the Gaussian elimination method. However, if $K < N$, there may be infinite number of solutions for $\hat{\mathbf{x}}$ as the coding coefficient matrix $\mathbf{C}$ is not full-rank. Hence, additional constraints need to be imposed into $\mathbf{C}$ appropriately such that the coding coefficient matrix becomes a full-rank matrix and the corresponding $\hat{\mathbf{x}}$ is a good approximation of $\mathbf{x}$. A similar problem has been studied in compressive sensing, where the original data can be recovered from a small set of equations, under sparsity assumptions [11]. However, such approaches are not applicable to this problem, since coding operations are performed in finite fields on data that are not necessarily sparse. When the data are correlated, the correlation information can be used by the decoder in order to provide additional constraints to $\mathbf{C}$ in the decoding process. This leads to approximate decoding solutions that enable the reconstruction of the original data with tolerable distortion. We study in the next section the influence of the finite field size (GF size) in the proposed framework, and then, we provide illustrative examples of approximate decoding of video and sensor data.

3. INFLUENCE OF FINITE FIELD SIZE

In this section, we study the impact of the construction of the coding coefficient matrix $\mathbf{C}$ on the approximated decoding performance. In particular, we analyze the influence of the GF size on the performance of the system. We assume that $\mathbf{C}$ is not a full-rank matrix. Approximate decoding is performed with help of additional constraints, which can be imposed based on the correlation information. The correlation information is communicated as side information in the beginning of transmission process.

We study first the influence of the GF size on the decoding error probability. Then, we determine the optimal GF size that leads to the smallest expected error at decoder.

Theorem 1 *Let x_n be selected from a finite size alphabet $\mathscr{X}$ and coding coefficients $c_n(k)$ be randomly selected from*

[2] In this paper, vectors and matrices are represented by boldfaced lowercase and boldfaced capital letters, respectively.

GF(2^r). If GF(2^r) is extended to GF(2^R) ($r < R$), the probability that decoding errors become higher increases.

Proof: Let $x \in \mathscr{X}$ be an original data, and let $\hat{x}_r$ and $\hat{x}_R$ be the decoded x over GF(2^r) and GF(2^R), respectively, where $R > r$ for $r, R \in \mathbb{N}$. We assume that $|\mathscr{X}| = 2^r$. We also assume that the recovered data is uniformly distributed over $\mathscr{X}$, since the coefficients of RLNC are randomly selected based on a uniform distribution over GF(2^r) or GF(2^R) in RLNC, the reconstructed data follows a uniform distribution [12]. Thus, the probability mass function of $\hat{x}_k$ is given by

$$p_k(\hat{x}_k) = \begin{cases} 1/2^k, & \text{if } x_k \in [0, 2^k - 1] \\ 0, & \text{otherwise} \end{cases}$$

for $k \in \{r, R\}$. To prove that extending GF size results in a higher decoding error, we may show

$$\Pr(|x - \hat{x}_R| \geq |x - \hat{x}_r|) > 0.5. \quad (2)$$

The left hand side of (2) can be expressed as

$$\begin{aligned} &\Pr\left(\hat{x}_R \geq \hat{x}_r, x \leq \frac{\hat{x}_R + \hat{x}_r}{2}\right) + \Pr\left(\hat{x}_R < \hat{x}_r, x > \frac{\hat{x}_R + \hat{x}_r}{2}\right) \\ &= \Pr(\hat{x}_R \geq \hat{x}_r) \Pr(2x \leq \hat{x}_R + \hat{x}_r | \hat{x}_R \geq \hat{x}_r) \\ &\quad + \Pr(\hat{x}_R < \hat{x}_r) \Pr(2x > \hat{x}_R + \hat{x}_r | \hat{x}_R < \hat{x}_r) \\ &= 2^{r-R-1} + \left(1 - 2^{r-R}\right) \hat{P} \end{aligned}$$

since $\hat{x}_R$ and $\hat{x}_r$ are both uniformly distributed. We define $\hat{P} \triangleq \Pr(2x \leq \hat{x}_R + \hat{x}_r | \hat{x}_R \geq \hat{x}_r)$. Using Bayes' rule,

$$\begin{aligned} \hat{P} &= \sum_{z=0}^{2^r-1} \Pr(2z \leq \hat{x}_R + \hat{x}_r | \hat{x}_R \geq \hat{x}_r, x = z) \Pr(x = z) \\ &= \frac{1}{2^r} \sum_{z=0}^{2^r-1} \Pr(2z \leq \hat{x}_R + \hat{x}_r | \hat{x}_R \geq \hat{x}_r, x = z). \end{aligned}$$

For $r, R \in \mathbb{N}$ and $R > r$, R can be expressed as $R = r + \alpha$, where $\alpha \in \mathbb{N}$, and since

$$\begin{aligned} &\sum_{z=0}^{2^r-1} \Pr(2z \leq \hat{x}_R + \hat{x}_r | \hat{x}_R \geq \hat{x}_r, x = z) \\ &= \frac{1}{2^{r+R}} \sum_{z=0}^{2^r-1} \left[2^{r+R} - \left\{2^{r-1}(2^r - 1) + 2\sum_{l=0}^{z} l\right\}\right] \\ &= \frac{1}{2^{r+R}} \left\{2^{2r+R} - \frac{1}{6}\left(5 \cdot 2^{3r} - 3 \cdot 2^{2r} - 2 \cdot 2^r\right)\right\}, \end{aligned}$$

$\hat{P}$ can be expressed as

$$\hat{P} = 1 - \frac{1}{6}\left[5 \cdot \frac{1}{2^\alpha} - \frac{3}{2^{r+\alpha}} - \frac{2}{2^{2r+\alpha}}\right].$$

Because $\lim_{r \to \infty} \hat{P} > 0.5$ for all $\alpha \in \mathbb{N}$, and $\hat{P}$ is a non-increasing function of r, $\hat{P} > 0.5$ for all r and R. Therefore,

$$\Pr(|x - \hat{x}_R| \geq |x - \hat{x}_r|) > 0.5 \quad (3)$$

which completes the proof. ■

Theorem 1 implies that a smaller GF size is preferred given a fixed number of data set, in order to reduce the expected decoding error. However, if the GF size becomes smaller, the maximum number of data that can be encoded and perfectly decoded by RLNC decreases correspondingly. Specifically, if $|\mathscr{X}| > 2^{r'}$ for $r' < r$, part of data in $\mathscr{X}$ needs to be discarded such that $|\mathscr{X}'| \leq 2^{r'}$. Hence, all the data in $\mathscr{X}'$ can be distinctly encoded in GF($2^{r'}$).

In summary, reducing the GF size for coding coefficients may result in lower decoding errors. However, this also induces higher loss of original data information. Based on this clear tradeoffs, Theorem 2 shows that

- there exists an optimal GF size that minimizes the expected decoding error,
- what is the optimal GF size.

In this analysis, we assume that if the GF size is reduced from GF(2^r) to GF(2^{r-z}), the least significant z bits are first discarded from $x \in \mathscr{X}$. Moreover, we assume that the corresponding information (data) loss is uniformly distributed and the recovered data is also uniformly distributed [12].

Theorem 2 *There exists an optimal GF size that minimizes the expected decoding error of RLNC encoded data. Moreover, the optimal GF size, GF(2^{r-z^*}), is determined at $z^* = \lceil (r-1)/2 \rceil$ and $z^* = \lfloor (r-1)/2 \rfloor$.*

Proof: Suppose that $|\mathscr{X}| = 2^r$ and GF(2^r). If GF size is reduced from GF(2^r) to GF(2^{r-z}), where $0 \leq z \leq r-1$ ($z \in \mathbb{Z}$), the decoding errors e_D are uniformly distributed over $[-r_D, r_D]$, where $r_D = 2^{r-1-z} - 1$, i.e.,

$$p_{e_D}(e_D) = \begin{cases} 1/(2r_D + 1), & \text{if } e_D \in [-r_D, r_D] \\ 0, & \text{otherwise} \end{cases}. \quad (4)$$

Correspondingly, $\mathscr{X}$ is reduced to $\mathscr{X}'$, where $|\mathscr{X}'| = 2^{r-z}$ by discarding least significant z bits from all $x \in \mathscr{X}$. This information loss also results in errors e_I over $[-r_I, r_I]$, where $r_I = 2^z - 1$, i.e.,

$$p_{e_I}(e_I) = \begin{cases} 1/(2r_I + 1), & \text{if } e_I \in [-r_I, r_I] \\ 0, & \text{otherwise} \end{cases}. \quad (5)$$

The distribution of total error, $p_{e_T}(e_T) = p_{e_D}(e_D) + p_{e_I}(e_I)$, is given by [13]

$$\begin{aligned} p_{e_T}(e_T) = \frac{1}{2} H\{&|e_T + r_I + r_D + 1| - |e_T + r_I - r_D| \\ &- |e_T - r_I + r_D| + |e_T - r_I - r_D - 1|\} \end{aligned}$$

for $|e_T| \leq r_I + r_D \triangleq e_T^{max}$ and $H = (2r_I + 1)^{-1}(2r_D + 1)^{-1}$. Since $e_T + r_I + r_D + 1 \geq 0$ and $e_T - r_I - r_D - 1 \leq 0$ for all $|e_T| \leq e_T^{max} (= r_I + r_D)$, by substituting r_I and r_D, we have

$$\begin{aligned} p_{e_T}(e_T) = \frac{1}{2} H\{&2\left(2^z + 2^{r-1-z} - 1\right) \\ &- |e_T + 2^z - 2^{r-1-z}| - |e_T - 2^z + 2^{r-1-z}|\}. \quad (6) \end{aligned}$$

By denoting $a(z) \triangleq 2^z - 2^{r-1-z}$ and $b(z) \triangleq 2^z + 2^{r-1-z}$, the expected decoding error $E[|e_T|] = \sum_{e_T=-\infty}^{\infty} |e_T| \cdot p_{e_T}(e_T)$ can be expressed as

$$\frac{1}{2} H \sum_{e_T = -e_T^{max}}^{e_T^{max}} |e_T| [2(b(z) - 1) - |e_T + a(z)| - |e_T - a(z)|].$$

Since both $|e_T|$ and $[2(b(z) - 1) - |e_T + a(z)| - |e_T - a(z)|]$ are symmetric on $z = \lceil (r-1)/2 \rceil$ and $z = \lfloor (r-1)/2 \rfloor$, $E[|e_T|]$ is also symmetric. Moreover, it can be easily shown that for

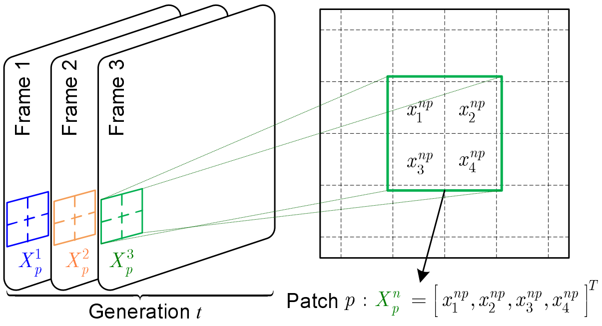

Figure 3: Illustrative examples of patch p with size 2×2 and patches in a generation t.

the case where $a(z) > 0$, which corresponds to $r/2 < z \leq r-1$, $E[e_T]$ can be expressed as

$$E[|e_T|] = \left[\frac{H}{3}b(z)(b(z)-1)(b(z)-2) - \frac{H}{3}a(z)(a(z)^2-1)\right]$$

and this is an increasing function for $r/2 < z \leq r-1$. Since $E[|e_T|]$ is a symmetric on $z = \lceil (r-1)/2 \rceil$ and $z = \lfloor (r-1)/2 \rfloor$, and is an increasing function over $r/2 < z \leq r-1$, $E[|e_T|]$ is convex over $0 \leq z \leq r-1$. Therefore, there exists an optimal z^* that minimizes the expected decoding error. Moreover, since $E[|e_T|]$ is symmetric on $\lceil (r-1)/2 \rceil$ and $\lfloor (r-1)/2 \rfloor$, the minimum $E[e_T]$ can be achieved at $z^* = \lceil (r-1)/2 \rceil$ and $z^* = \lfloor (r-1)/2 \rfloor$. ■

In the next section, we confirm the theoretical results discussed in this section by implementing the approximate decoding strategy to two illustrative applications.

4. ILLUSTRATIVE EXAMPLES

4.1 Network Coding of Uncompressed Video Frames

We illustrate the findings of the previous section in an application that perform network coding of images in a video sequence, and approximate decoding using the information provided by motion estimation in the video sequence.

Let $Y_p^t(k) = \sum_{n=1}^{N} c_p^n(k) X_p^n$ be the kth received data that corresponds to a patch p with size $L \times L$ in generation t (i.e., a GOP). $Y_p^t(k) = [y_1^{tp}(k), \ldots, y_{L^2}^{tp}(k)]^T$ is a vector of linear combination of corresponding $X_p^n = [x_1^{np}, \ldots, x_{L^2}^{np}]^T$ at nth frame in GOP t with coding coefficients $c_p^n(k)$, which is randomly chosen in GF(2^r). We assume that the original data (i.e., pixels) has values ranging in $[0, 255]$, and thus, $|\mathscr{X}| = 256 = 2^8$. An example is illustrated in Fig. 3.

For patch p, if a node receives K innovative data, i.e., the node has $Y_p^t(k)$, $k = 1, \ldots, K$, the node can form the following linear system $\mathbf{Y}_p^t(K) = \mathbf{C}_p \mathbf{X}_p$:

$$\begin{bmatrix} Y_p^t(1) \\ \vdots \\ Y_p^t(K) \end{bmatrix} = \begin{bmatrix} c_p^1(1) I_{L^2} & \cdots & c_p^N(1) I_{L^2} \\ \vdots & \ddots & \vdots \\ c_p^1(K) I_{L^2} & \cdots & c_p^N(K) I_{L^2} \end{bmatrix} \begin{bmatrix} X_p^1 \\ \vdots \\ X_p^N \end{bmatrix},$$

where I_{L^2} is $L \times L$ identity matrix. Correspondingly, for N frames (each of frame is decomposed into M patches, $1 \leq p \leq M$), linear system $\mathbf{Y}(K) = \mathbf{C}\mathbf{X}$ in generation t can also be formed, where $\mathbf{Y}(K) = [\mathbf{Y}_1^t(K), \ldots, \mathbf{Y}_M^t(K)]^T$, $\mathbf{C} = diag(\mathbf{C}_1, \ldots, \mathbf{C}_M)$, and $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_M]^T$. $\mathbf{Y}(K)$ is a $KML^2 \times 1$ vector, $\mathbf{C}$ is a $KML^2 \times NML^2$ matrix, and $\mathbf{X}$ is a $NML^2 \times 1$ vector. We consider the case of $K < N$, where the approximated decoding approach can be deployed by imposing additional $(N-K)ML^2$ constraints into the coding coefficient matrix $\mathbf{C}$.

The illustrative example consists of the first three frames extracted from *Silent* QCIF format (174×144) standard sequence. The pixel values in each frame have range of [0, 255]. Thus, RLNC coefficients are randomly selected from GF(2^8), which corresponds to $r = 8$ in our analysis of Section 3. However, if the GF size decreases by z bits, i.e., the network coding coefficients are selected over GF(2^{8-z}), the least significant z bits are discarded from each pixel. In our experiments, we assume that $2/3$ of innovative packets are received. The rest of equations can be additionally imposed based on the information about the matched units in a patch (i.e., correlations between frames). To find the matched units in a patch a simple block-based motion estimation technique is used.[3] Each of the constraints can contribute to $\mathbf{C}$ as a form of row vector with NML^2 zeros (i.e., additive identity over GF(2^r)) except two elements of 1 and -1 in the position of matched units.[4] For the motion estimation, we use the first and the second frames. The patch size is 16×16 and a block size is 8×8. The experiment results are shown in Fig. 4.

Fig. 4 shows the quality measured as PSNR from actually decoded the three frames in *Silent* for different GF sizes 2^{8-z}. As discussed in Theorem 2, the expected decoding error can be minimized if $z^* = \lceil (r-1)/2 \rceil$ or $z^* = \lfloor (r-1)/2 \rfloor$, which corresponds to $z^* = 3$ and $z^* = 4$. This is confirmed from the experiment results, where the two highest average PSNRs are achieved at $z = 3$ and $z = 4$.

4.2 Seismic Signals in Sensor Networks

The approximated decoding can be used to recover the data transmitted from distributed sensors in sensor networks, where each sensor captures a source signal from different locations. The correlation among the signals that depends on the proximity of the sensors used for approximate decoding when the decoding system is not full-rank. Note that the closer the sensors are located, the higher correlations are achieved.

We consider seismic signals that are captured by sensors spaced by 100m transmitted to neighbors nodes (e.g., relay nodes) or receivers. For this illustration, we assume that a receiver tries to recover the signals of sensors 1, 2, and 30 from the received packets in a generation (i.e., signal samples in a fixed size window) encoded based on RLNC. We assume that the correlation information among signals is included in addition to the coding coefficients in the encoding process, and transmitted with packets. Similarly to the illustrative example in Section 4.1, we assume that $2/3$ of linear equations required for perfect decoding are received, and that the rest of $1/3$ of constraints are imposed into the coding coefficient matrix based on the included correlation information. For simplicity, we assume that the signals from two close sensors, sensor 1 and sensor 2, are highly correlated. The cap-

[3]This information is included in the encoded packets in addition to the coding coefficients.

[4]-1 denotes the additive inverse of 1 over GF(2^r).

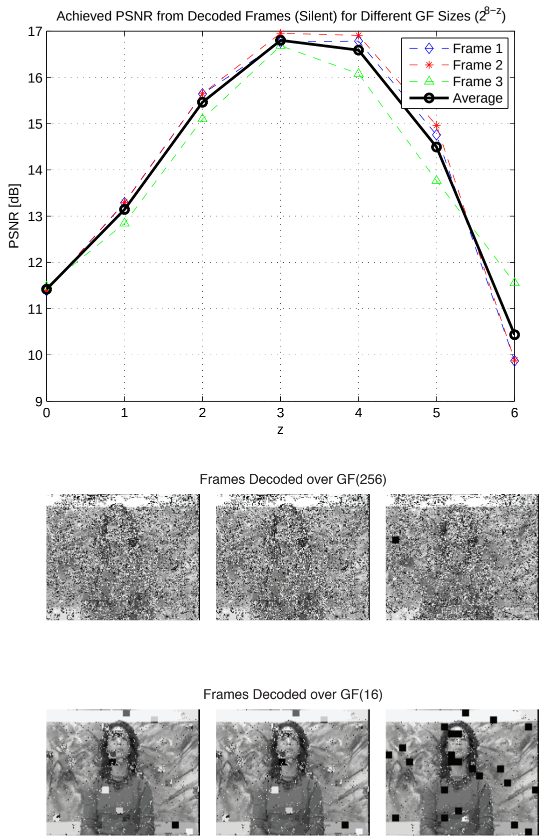

Figure 4: Achieved PSNR for different GF sizes (i.e., GF(2^{8-z})) and correspondingly recovered frames (*Silent*).

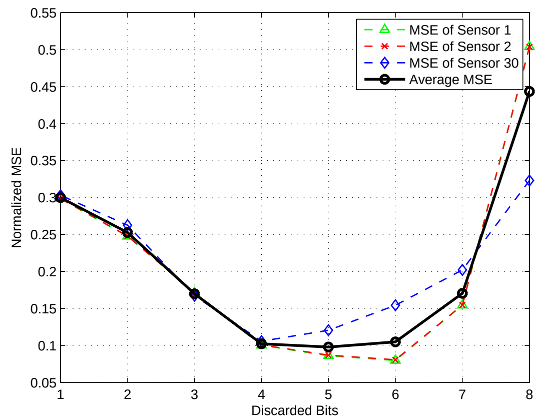

Figure 5: Normalized average mean square error (MSE) for different GF sizes (i.e., GF(2^{10-z}))

tured data is in the range of $[0, 1023]$. Thus, the maximum GF size is GF(2^{10}). The experimental results are shown in Fig. 5.

Fig. 5 shows the mean square error (MSE) from decoded signals for different GF sizes sizes 2^{10-z}. Theorem 2 is also confirmed from these results, as the two lowest average decoding errors (i.e., normalized average MSE) are achieved at $z^* = \lceil(10-1)/2\rceil = 5$ and $z^* = \lfloor(10-1)/2\rfloor = 4$.

5. CONCLUSIONS

In this paper, we have described a framework for the delivery of correlated information sources with help of network coding and approximate decoding based on correlation information. We have analyzed the tradeoffs between the decoding performance and the size of finite fields. We can determine an optimal field size that leads to the highest approximated decoding performance. The proposed approach is implemented in illustrative video streaming and sensor networks applications, where the experimental results confirm the effectiveness of the proposed approach. Further interesting research topics may include the relation between delay and approximate error

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[2] P. A. Chou and Yunnan Wu, "Network coding for the internet and wireless networks," *IEEE Signal Processing Mag.*, vol. 24, no. 5, pp. 77–85, Sept. 2007.

[3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[4] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in *Proc. Allerton Annual Conf. Commun., Control, and Comput.*, Monticello, IL, USA, Oct. 2003.

[5] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Commun., Control and Comput.*, Monticell, IL, USA, Oct. 2003.

[6] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[7] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 155–170, Jan. 2006.

[8] T. P. Coleman, E. Martinian, and E. Ordentlich, "Joint source-channel coding for transmitting correlated sources over broadcast networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 8, pp. 3864–3868, Aug. 2009.

[9] M. Yang and Y. Yang, "A linear inter-session network coding scheme for multicast," in *Proc. IEEE Int. Symp. Netw. Comput. and Applications*, 2008, pp. 177–184.

[10] A. Neumaier, "Solving ill-conditioned and singular linear systems: A tutorial on regularization," *SIAM Review*, vol. 40, no. 3, pp. 636 –666, Sept. 1998.

[11] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[12] D. Bogdanov, S. Laur, and J. Willemson, "Sharemind: A framework for fast privacy-preserving computations," *Springer Lecture Notes in Computer Science*, vol. 2583/2008, pp. 192–206, Oct 2008.

[13] D. M. Bradley and R. C. Gupta, "On the distribution of the sum of *n* non-identically distributed uniform random variables," *Journal Annals of the Institute of Statistical Mathematics*, vol. 54, no. 3, pp. 689 –700, Sept. 2002.