Opportunistic Noisy Network Coding for Fading Relay Networks Without CSIT

Sang-Woon Jeon, *Member, IEEE*, Sung Hoon Lim, *Member, IEEE*, and Bang Chul Jung, *Senior Member, IEEE*

***Abstract*—The parallel relay network is studied, in which a single source node sends a message to a single destination node with the help of *N* parallel relays. Channel coefficients are assumed to vary over time and channel state information (CSI) is causally available only at the receiver side (CSIR). Opportunistic noisy network coding is proposed for intelligently exploiting CSIR at each relay in a distributed manner by operating the noisy network coding scheme with adaptive compression. More specifically, each relay opportunistically vector-quantizes the collection of received symbols that is received with channel gains larger than a certain threshold. It then forwards the digital compression information to the destination node using independently generated Gaussian codes. For independent and identically distributed (i.i.d.) Rayleigh fading, the proposed scheme is shown to achieve the ergodic capacity in the large number of relays regime. Furthermore, the proposed scheme is extensively compared with several alternative schemes, the decode-forward scheme, the adaptive amplify-forward scheme, and the non-adaptive noisy network coding scheme over geometric models. We show that the new proposed scheme provides significant gain over these schemes in various cases.**

***Index Terms*—Adaptive compression, approximate capacity, compress-forward, fast fading, noisy network coding, parallel relay network.**

I. Introduction

In recent years, cooperative communication using relays has been considered as a promising technique to improve the spectral efficiency and coverage of wireless networks. For such systems, the fundamental design principles for optimal relaying has been the primary concern. A canonical model capturing this feature is the *parallel relay network* [1]. The parallel relay network is a two hop network in which the source node communicates to the destination node by the help of *N* relay nodes. The source node transmits to a set of relays through a broadcast channel, and the relay nodes transmits to the destination node through a multiple access channel as depicted in Fig. 1.

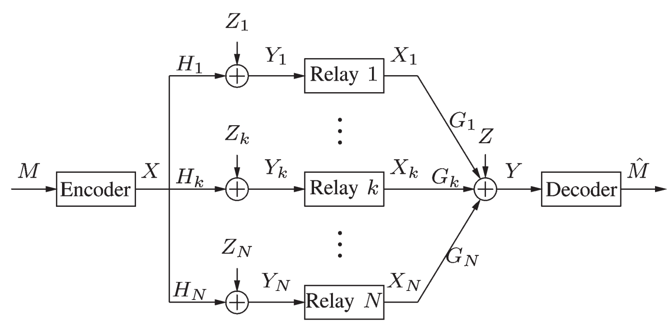

Fig. 1. The *N* relay fading parallel relay network.

For Gaussian relay networks, there exist three fundamental relaying strategies: *decode-forward (DF), compress-forward (CF), and amplify-forward (AF)*. In the DF scheme, originally proposed by Cover and El Gamal [2], the relay recovers the message of the source either fully or partially and forwards it to the destination by coherent transmission. The DF scheme has been extended to networks with arbitrary topology by Xie and Kumar [3] and Kramer, Gastpar, and Gupta [4], in which every relay node along the path from the source to the destination decodes and forwards the message. In the CF scheme, again proposed by Cover and El Gamal [2], the relay instead sends a description of its noisy observation by first compressing the observation and forwarding the compression information to the destination. Due to its simplicity, the relay operation in the CF scheme is less sensitive to the end-to-end operations at the source and destination, making it more attractive than DF for large scale networks [4]. The CF scheme has been extended to networks with arbitrary topology by Lim, Kim, El Gamal, and Chung [5] and Yassaee and Aref [6] in the context of noisy network coding. Further extensions of noisy network coding using short message block Markov encoding has been proposed in [7], [8]. Specializing the noisy network coding scheme for Gaussian networks with *N* nodes, it was shown in [5] that noisy network coding is universally within 1.26*N* bits/s/Hz of the capacity, which refines upon the previously established gap in [9]. Alternatively, the AF scheme is another relaying paradigm widely considered specifically for Gaussian relay networks [1],

Manuscript received November 27, 2014; revised April 7, 2015; accepted June 5, 2015. Date of publication June 30, 2015; date of current version November 9, 2015. The work of S.-W. Jeon was supported in part by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (MEST) (NRF-2013R1A1A1064955). The work of S.H. Lim was supported in part by the European ERC Starting Grant 259530-ComCom. The work of B.C. Jung was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2A10004905). The material in this paper was presented in part at the IEEE Globecom, Houston, TX, USA, December 2011. The associate editor coordinating the review of this paper and approving it for publication was Z. Han.

S.-W. Jeon is with the Department of Information and Communication Engineering, Andong National University, Andong 760-749, Korea (e-mail: swjeon@anu.ac.kr).

S. H. Lim is with the School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne 1015, Switzerland (e-mail: sung.lim@epfl.ch).

B. C. Jung is with the Department of Electronics Engineering, Chungnam National University, Daejon 305-764, Korea (e-mail: bangchuljung@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2448085

[10]–[15]. In the AF scheme, the relay simply sends a scaled version of its received signal within the relay power constraint.

For the Gaussian parallel relay network with $N = 2$, the achievable rates of DF and AF have been analyzed in [16] showing that DF and AF achieve the capacity in some signal-to-noise ratio (SNR) regimes. It was further shown in [9] that partial DF can achieve capacity to within 1 bit/sec/Hz, independent of SNR and channel parameters. When the number of relays N tends to infinity, it was shown that AF can achieve the capacity in certain SNR regimes. The case when there is a bandwidth mismatch between the first and second hop was studied in [17], [18]. More recently, it was shown in [15] that the bursty AF scheme achieves the capacity of the symmetric Gaussian parallel relay network to within a constant gap independent of SNR and the number of relays N.

Motivated by the approximate (finite gap) capacity results for Gaussian parallel relay networks, we further investigate the capacity characteristics of *fading* relay networks with increasing number of relays. In particular, the fast fading model with CSI only at the receiver side (CSIR) is considered, which makes our work distinguishable from other models that assume block fading or global CSI at all nodes [19]–[24]. We propose the *opportunistic* noisy network coding scheme as a novel extension of the noisy network coding scheme, in which the relay observation is opportunistically compressed by adapting on the source-relay CSI at the relay node. Each relay node vector quantizes its observation sequence adaptively based on the CSI information. Conceptually, the proposed vector quantization scheme effectively compresses the subset of the symbols with channel gains above a certain threshold, while the rest of the symbols are simply neglected. Then, the relays send the digital compression information to the destination node using independently generated Gaussian codes. We show that this simple threshold-based adaptation scheme achieves the capacity in the large number of relays regime while strictly outperforming other schemes such as AF and DF. Unlike the conventional opportunistic or adaptive approaches relying on CSIT, our work demonstrates how adaptation based on *CSIR* can be beneficial.

The rest of the paper is organized as follows. We begin with a formal statement of our problem in Section II. We first present our main results in Section III. In Section IV, the detailed description of the opportunistic noisy network coding scheme and the proof of achievability is given. In Section V, we compare the opportunistic noisy network coding scheme with various schemes. We briefly discuss generalizations to more general fading distributions and network configurations in Section VI. Finally, Section VII concludes the paper.

Throughout the rest of the paper, we adopt the notation in [25]. In particular, otherwise specifically stated, we denote random variables with upper-case letters and denote their realizations with the corresponding lower-case letters. The expectation of a random variable A is denoted by $\mathsf{E}(A)$ and $\mathsf{C}(x) = \log(1+x)$, where the log operation is with respect to base two. For set notation we use calligraphic letters, e.g., $\mathcal{S}$, and denote $[1:N] = \{1, 2, \cdots, N\}$. For a subset $\mathcal{S} \subseteq [1:N]$ its complementary set is represented by $\mathcal{S}^c = [1:N] \setminus \mathcal{S}$ and the cardinality of a set is represented by $|\mathcal{S}|$. We also use the notation $A(\mathcal{S}) = \{A_k, k \in \mathcal{S}\}$ and $A^N = \{A_1, A_2, \cdots, A_N\}$.

II. Problem Statement

A. Fading Parallel Relay Networks

We consider the fading parallel relay network depicted in Fig. 1 in which the source node wishes to send a message to the destination node with the help of N relay nodes. The source node has a channel input X, relay node $k \in [1:N]$ has a channel input X_k and observes a channel output Y_k, and the destination node observes a channel output Y. The input-output relations at time t are given by

$$Y_k[t] = H_k[t]X[t] + Z_k[t], \qquad k \in [1:N] \tag{1}$$

and

$$Y[t] = \sum_{k=1}^{N} G_k[t]X_k[t] + Z[t] \tag{2}$$

where $Z_k[t]$, $k \in [1:N]$ and $Z[t]$ are independent complex Gaussian noise with $\mathcal{N}_{\mathbb{C}}(0, 1)$. We assume time varying channels such that at time t, $H_k[t]$ is drawn from $\mathcal{N}_{\mathbb{C}}(0, \sigma^2_{H_k})$ which is assumed to be independent of other channel coefficients from different links and time indices. Similarly, $G_k[t]$ is drawn from $\mathcal{N}_{\mathbb{C}}(0, \sigma^2_{G_k})$ and is assumed to be independent of other channel coefficients from different links and time indices. We further assume that CSI is available only at receiver sides, i.e., at the end of n transmissions, relay node k knows H_k^n and the destination knows H_1^n to H_N^n and G_1^n to G_N^n. We assume average power constraint P for the source node and P_r/N for each relay node, i.e., $\mathsf{E}[|X[t]|^2] \le P$ and $\mathsf{E}[|X_k[t]|^2] \le P_r/N$ for all $k \in [1:N]$. Hence, the total transmit power of all relay nodes is upper bounded by P_r.

We would like to emphasise two properties regarding our setup. First, by the freedom of the choice of the channel gain variances and the power constraints, the setup is without loss of generality in the sense that it can cover all possible received *SNR* settings. Second, for any number of relay nodes, we normalise the total amount of power that the relays are allowed to consume by P_r, however, this does not mean that all the relays share power among each other, but rather we assume that each relay has an individual power constraint P_r/N. The main motivation for this assumption is that we wish to focus on the “opportunistic gain” that the N relay nodes provide by discarding the effect of increased power from having more relay nodes. In the rest of the paper, we will frequently omit the time index for notational convenience.

Let $[1:2^{nR}]$ be the message set of the source. A $(2^{nR}, n)$ code consists of an encoding function $x^n(m)$, $m \in [1:2^{nR}]$, which maps a message m into a length-n input sequence, relay encoding functions $x_k[t] = \varphi_{k,t}(y_k^{t-1}, h_k^{t-1})$, for $t \in [1:n]$ and $k \in [1:N]$, which at time t maps a length-$(t-1)$ output sequence and a length-$(t-1)$ CSI sequence into an input symbol, and a decoding function $\hat{m}(y^n, h_1^n, \cdots, h_N^n, g_1^n, \cdots, g_N^n) \in [1:2^{nR}]$, which maps a length-$n$ output sequence and a set of length-n CSI sequences to a message estimate. We assume that the message M is uniformly distributed over $[1:2^{nR}]$ and define the average probability of error as $\mathsf{P}\{\hat{M} \neq M\}$. A rate R is said to be *achievable* if there exist a sequence of $(2^{nR}, n)$ codes with $\mathsf{P}\{\hat{M} \neq M\} \to 0$ as $n \to \infty$. The capacity C_N of the fading parallel relay network with N relay nodes is the supremum of all achievable rates. When the context is clear, we will drop the subindex N throughout the paper.

III. Main Results

In this section, we first state our main result which establishes a lower bound on the capacity of fading parallel relay networks. The lower bound is attained by the *opportunistic noisy network coding* (ONNC) scheme. The opportunistic noisy network coding scheme is presented in two steps. First, we present the noisy network coding scheme with short messages and block Markov encoding, similar to the short message noisy network coding schemes presented in [6]–[8]. However, different form the previous approaches, the decoder recovers each message with only a one block delay without binning at the relay nodes. The details are explain in Section IV and Appendix A. The resulting opportunistic noisy network coding achievable rate is given by:

$$C \geq \max \min_{\mathcal{S}\subseteq[1:N]} \Big\{ I\left(X; \hat{Y}(\mathcal{S})|H^N\right) + I\left(X(\mathcal{S}); Y|X(\mathcal{S}^c), G^N\right) - I\left(\hat{Y}(\mathcal{S}); Y(\mathcal{S})|X, \hat{Y}(\mathcal{S}^c), H^N\right)\Big\} \quad (3)$$

where the maximization is over all probability distributions $p(x)\prod_{k=1}^{N} p(x_k)p(\hat{y}_k|y_k, h_k)$ such that the power constraints are satisfied.

The rate expression in (3) involves a maximization step over $p(\hat{y}_k|y_k, h_k)$. Maximising over $p(\hat{y}_k|y_k, h_k)$ can be interpreted as choosing a good vector quantizer, i.e., we find a sequence $\hat{y}_k^n$ that is jointly typical (with respect to the chosen distribution) with the observations (y_k^n, h_k^n). Accordingly, in the second step, we provide a heuristic adaptive compression scheme. In this new approach, each relay node adaptively compresses its observation based on its received CSI (i.e., H_k^n) instead of using fixed compression rates as done for Gaussian networks. The details of the compression scheme is explained in Section IV.

To present our main result, consider a real number $\alpha_k \in (0, 1]$, $k \in [1 : N]$ such that $\mathsf{P}\{|H_k|^2 \geq \gamma_k\} = \alpha_k$. Equivalently, we have $\gamma_k = \sigma_{H_k}^2 \ln(1/\alpha_k)$. Define $\tilde{H}_k$ as the truncated random variable of H_k conditioned on $|H_k|^2 \geq \gamma_k$, in which the probability distribution of $|\tilde{H}_k|^2$ is given by

$$p_{|\tilde{H}_k|^2}(x) = \begin{cases} p_{|H_k|^2}(x)/\alpha_k & \text{if } x \geq \gamma_k, \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $p_{|H_k|^2}(x) = \frac{1}{\sigma_{H_k}^2} e^{-x/\sigma_{H_k}^2}$. We are ready to state our main theorem.

Theorem 1 (ONNC Lower Bound): For the fading parallel relay network, the capacity is lower bounded as

$$C \geq \max \min_{\mathcal{S}\subseteq[1:N]} R_{\mathrm{ONNC}}(\mathcal{S}) \quad (5)$$

where

$$\begin{aligned} R_{\mathrm{ONNC}}(\mathcal{S}) &= \sum_{\Lambda\subseteq\mathcal{S}^c} \left(\prod_{j\in\mathcal{S}^c\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right) \\ &\quad \times \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c\setminus\Lambda} \frac{|\tilde{H}_k|^2 P}{1+Q_k}\right)\right] + \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}} |G_k|^2 \frac{P_r}{N}\right)\right] \\ &\quad - \sum_{\Lambda\subseteq\mathcal{S}} \left(\prod_{j\in\mathcal{S}\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right) \sum_{k\in\mathcal{S}\setminus\Lambda} \mathsf{C}\left(\frac{1}{Q_k}\right) \end{aligned} \quad (6)$$

and the maximization is taken over all $\alpha_k \in (0, 1]$ and $Q_k > 0$, $k \in [1 : N]$.

The detailed operations of the opportunistic noisy network coding scheme as well as the proof of Theorem 1 is presented in Section IV.

To best demonstrate the performance of opportunistic noisy network coding, we compare its performance with the cut-set upper bound. The cut-set upper bound on the capacity C of the fading parallel relay network simplifies to

$$C \leq \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c} |H_k|^2 P\right) + \mathsf{C}\left(\sum_{k\in\mathcal{S}} |G_k|^2 \frac{P_r}{N}\right)\right]. \quad (7)$$

The key observation here is that independent Gaussian inputs at each node simultaneously maximize every cut under time-varying channel coefficients without CSIT. We refer to Appendix B for the proof. By comparing the upper and lower bounds, we have the following theorem.

Theorem 2 (Asymptotic Capacity for Symmetric Fading): For the fading parallel relay network with $\sigma_{H_k} = \sigma_{G_k} = 1$,

$$\lim_{N\to\infty} C_N = \mathsf{C}(P_r) \quad (8)$$

for any P and P_r.

The proof of the theorem is given in Section IV. The theorem implies that the opportunistic noisy network coding scheme achieves the capacity of parallel relay networks for the symmetric fading case, universally for any P and P_r as the number of relays N increases. Notice that, since our model considers arbitrary P and P_r, the symmetry assumption $\sigma_{H_k} = \sigma_{G_k} = 1$ considers every case where $\sigma_{H_k} = \sigma_H$ and $\sigma_{G_k} = \sigma_G$ for all k ($\sigma_H \neq \sigma_G$ in general). Furthermore, this asymptotic optimality is a rare property in that AF and DF cannot attain Theorem 2 (even by applying similar adaptations). A detailed comparison between the achievable rate of the opportunistic noisy network coding scheme with those of the AF and DF schemes is given in Section V-C.

IV. Proof of Achievability

In this section, we present the opportunistic noisy network coding scheme. The noisy network coding lower bound [5] for discrete memoryless networks can be adapted for the fading parallel relay network with power constraint and state dependency, i.e., random channel gains. Further taking advantage of the Markov structure of the network, the noisy network coding lower bound for fading parallel relay networks yields the lower bound (3). The achievability of (3) follows directly from the noisy network coding lower bound by treating (Y_k, H_k), $k \in [1:N]$ as the relay outputs and (Y, H^N, G^N) as the destination output of a discrete memoryless network, i.e.,

$$p(y^N, y, h^N, g^N | x, x^N) = p(h^N)\, p(g^N)\, p(y^N, y \,|\, x, x^N, h^N, g^N). \tag{9}$$

However, by taking advantage of the (topologically) simple network structure, the general purpose noisy network coding scheme can be simplified in many ways. For completeness, in the following section, we provide the opportunistic noisy network coding scheme with modified random coding steps specifically tailored for parallel fading Gaussian networks and highlight the difference in choosing the compression codes. The modified scheme uses short message block Markov encoding as in [6]–[8]. The differences from these schemes are, first, we incorporate state (fading) dependency into our scheme and, second, by specially tailoring the noisy network coding scheme to the layered structure of the network, we have a simpler strategy. In particular, we propose a one-block delay forward decoder which is different from sliding window decoding [4], while the relays do not use binning.

A. Opportunistic Noisy Network Coding

We provide the proof for discrete memoryless networks. The extension to the Gaussian network is a straight forward extension of the quantization method given in [25]. We use a block Markov coding scheme in which a sequence of b i.i.d. messages $m_j \in [1:2^{nR}], j \in [1:b]$, is sent over $b+1$ blocks each consisting of n transmissions. The overall transmission rate is thus, $\frac{bR}{b+1}$, which tends to R as $b \to \infty$.

Codebook Generation: Fix $p(x)\prod_{k=1}^{N} p(x_k)p(\hat{y}_k|y_k, h_k)$. For the source node, for $j = 1, \ldots, b$, randomly and independently generate 2^{nR} sequences $x_j^n(m_j)$, $m_j \in [1:2^{nR}]$, according to $\prod_{i=1}^{n} p_X(x_i)$. Similarly, for $k = 1, \ldots, N$ and $j = 1, \ldots, b+1$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $x_{kj}^n(l_{k,j-1})$, $l_{k,j-1} \in [1:2^{n\hat{R}_k}]$, each according to $\prod_{i=1}^{n} p_{X_k}(x_{ki})$. For each $k = 1, \ldots, N$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $\hat{y}_{kj}^n(l_{kj})$, $l_{kj} \in [1:2^{n\hat{R}_k}]$, each according to $\prod_{i=1}^{n} p_{\hat{Y}_k}(\hat{y}_{ki})$. This defines the codebook

$$\mathcal{C}_j = \Big\{ x_j^n(m_j), x_{kj}^n(l_{k,j-1}), \hat{y}_{kj}^n(l_{kj}) : \\ m_j \in [1:2^{nR}], l_{kj}, l_{k,j-1} \in [1:2^{n\hat{R}_k}] \text{ for } k \in [1:N] \Big\},$$

where $j \in [1:b]$.

Encoding: To send the message m_j at block $j \in [1:b]$, the codeword $x_j^n(m_j)$ is transmitted.

Relay Encoding (Vector Quantization): At relay k, upon receiving y_{kj}^n at the end of block $j \in [1:b]$, it finds an index l_{kj} such that

$$\left(\hat{y}_{kj}^n(l_{kj}), y_{kj}^n, h_{kj}^n\right) \in \mathcal{T}_{\epsilon'}^{(n)}.$$

If there is more than one such index, choose one of them at random. If there is no such index, choose an arbitrary index at random from $[1:2^{n\hat{R}_k}]$. The codeword $x_{k,j+1}^n(l_{kj})$ is transmitted in the next block.

Decoding: Let $\epsilon > \epsilon' > 0$. At the end of block $j+1$, the decoder finds the unique message $\hat{m}_j \in [1:2^{nR}]$ such that

$$\left(x_j^n(\hat{m}_j), \hat{y}_{1j}^n(\hat{l}_{1j}), \ldots, \hat{y}_{Nj}^n(\hat{l}_{Nj}), h_{1j}^n, \ldots, h_{Nj}^n\right) \in \mathcal{T}_{\epsilon}^{(n)},$$

and

$$\Big(x_{1,j+1}^n(\hat{l}_{1j}), \ldots, x_{N,j+1}^n(\hat{l}_{Nj}), y_{j+1}^n, \\ h_{1,j+1}^n, \ldots, h_{N,j+1}^n, g_{1,j+1}^n, \ldots, g_{N,j+1}^n\Big) \in \mathcal{T}_{\epsilon}^{(n)}$$

for some $\hat{l}_{1j}, \ldots, \hat{l}_{Nj}$. If there is none or more than one such message, it declares an error.

Remark 1: Note that this decoder has two simultaneous joint typicality tests on sequences that belong to two consecutive block transmissions. The first joint typicality condition is on block j with the codewords and observation sequences that correspond to the communication of the first hop (broadcast), and the second joint typicality condition is on block $j+1$ with the codewords and observation sequences that correspond to the communication of the second hop (multiple access).

In Appendix A, we show that the probability of decoding error tends to zero as $n \to \infty$ if

$$\hat{R}_k > I(\hat{Y}_k; Y_k|H_k) + \delta(\epsilon'), \tag{10}$$

$$R + \hat{R}(\mathcal{S}) < I\left(X; \hat{Y}(\mathcal{S})|H^N\right) + I\left(X(\mathcal{S}); Y|X(\mathcal{S}^c), G^N\right) \\ + \sum_{k\in\mathcal{S}} I\left(\hat{Y}_k; X, \hat{Y}(\mathcal{S}_k), \hat{Y}(\mathcal{S}^c), H^N\right) - \delta(\epsilon), \tag{11}$$

where $\mathcal{S}_k = \{\mathcal{S} \cap [1:k-1]\}$ for all $\mathcal{S} \subseteq [1:N]$. Finally, by eliminating $\hat{R}_k$ using Fourier-Motzkin elimination, equations (10) and (11) simplify to

$$R < I\left(X; \hat{Y}(\mathcal{S})|H^N\right) + I\left(X(\mathcal{S}); Y|X(\mathcal{S}^c), G^N\right) \\ - I\left(\hat{Y}(\mathcal{S}); Y(\mathcal{S})|X, \hat{Y}(\mathcal{S}^c), H^N\right) - \delta(\epsilon) \tag{12}$$

for all $\mathcal{S} \subseteq [1:N]$.

For Gaussian networks, we choose the distribution $p(x)\prod_{k=1}^{N} p(x_k)p(\hat{y}_k|y_k, h_k)$ such that $X \sim \mathcal{N}_{\mathbb{C}}(0, P)$, $X_k \sim \mathcal{N}_{\mathbb{C}}(0, P_r/N)$, and $\hat{Y}_k = Y_k + \hat{Z}_k$ where $\hat{Z}_k \sim \mathcal{N}_{\mathbb{C}}(0, \eta_k(h_k))$ for $k \in [1:N]$. The function $\eta_k(\cdot) > 0$ is an arbitrary function of h_k that will be defined later in the section. By evaluation using the distribution above, we establish the following lower bound on the capacity,

$$C \geq \max \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\Bigg[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c} \frac{|H_k|^2 P}{1+\eta_k(H_k)}\right) \\ + \mathsf{C}\left(\sum_{k\in\mathcal{S}} |G_k|^2 \frac{P_r}{N}\right) - \sum_{k\in\mathcal{S}} \mathsf{C}\left(\frac{1}{\eta_k(H_k)}\right)\Bigg] \tag{13}$$

where the maximization is taken over all functions $\eta_k(h_k) > 0$, $k \in [1:N]$.

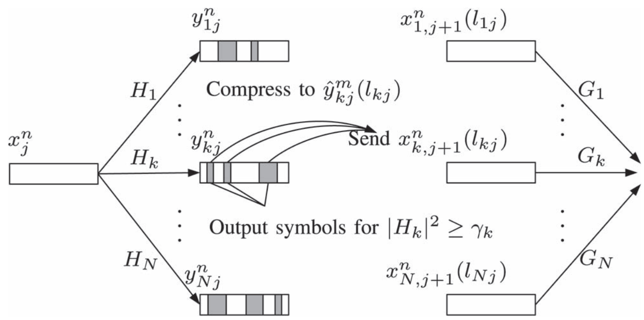

Fig. 2. Conceptual illustration of the threshold-based adaptation for the opportunistic noisy network coding scheme.

The achievable rate expression requires optimization over all functions $\eta_k(h_k)$, which itself is intractable for most cases. In our opportunistic noisy network coding scheme, we propose a *threshold*-based adaptation function, by choosing $\eta_k(h_k)$ as

$$\eta_k(h_k) = \begin{cases} Q_k & \text{if } |h_k|^2 \geq \gamma_k, \\ \infty & \text{otherwise} \end{cases} \tag{14}$$

where $Q_k > 0$. Intuitively, the proposed function $\eta_k(h_k)$ can be interpreted as having γ_k as a threshold for opportunistic compression in which relay k only compresses the observation symbols that is received with channel gains above this threshold with compression noise variance Q_k. To see this, recall that $\hat{y}^n(l_k)$ is a *vector quantization* of the pair (y_k^n, h_k^n) such that it holds the property that they are jointly typical with respect to the joint distribution $p(y_k, h_k)p(\hat{y}_k|h_k, y_k)$ where $p(\hat{y}_k|h_k, y_k) = \mathcal{N}_{\mathbb{C}}(y_k, \eta_k(h_k))$. This choice can be interpreted as quantizing the sub-sequence of y^n with $h_k > \gamma_k$ within distortion Q_k while the rest of the sequence has infinite distortion. Since the decoder has CSIR, it knows the location of which sequence is quantized within which level.

Fig. 2 gives a conceptual illustration on how the threshold-based adaptation operates in the opportunistic noisy network coding scheme. For relay node k at block j, the collection of outputs with channel gains above γ_k is compressed to $\hat{y}_{kj}^m(l_{kj})$, where $m \leq n$ is the number of symbols with $|H_k|^2 \geq \gamma_k$. While this compression step is a n length joint typical encoding based compression scheme, the index l_{kj} contains no information of the symbols that have channel gains below γ_k. The compression index l_{kj} is then sent by independently generated Gaussian codes $x_{k,j+1}^n(l_{kj})$ in the next block. Accordingly, the total n-length transmission at the relay is used in sending the index l_{kj}, which carries the compression information of a subset of the observation symbol y_k^n. As a result, the outputs with channel gains higher than the threshold are opportunistically compressed and forwarded to the destination.

We are now ready to prove Theorem 1. Consider the threshold-based adaptation function in (14). Recall $\alpha_k \in (0, 1]$ such that $\mathsf{P}\{|H_k|^2 \geq \gamma_k\} = \alpha_k$. By applying (14) in (13), we have

$$C \geq \max \min_{\mathcal{S}\subseteq[1:N]} \sum_{\Lambda\subseteq[1:N]} \left(\prod_{j\in[1:N]\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right) \times \mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda) + \mathsf{C}_2(\mathcal{S}) - \mathsf{C}_3(\mathcal{S}\setminus\Lambda)\right], \tag{15}$$

where

$$\mathsf{C}_1(\mathcal{S}) = \mathsf{C}\left(\sum_{k\in\mathcal{S}} \frac{|\tilde{H}_k|^2 P}{1+Q_k}\right), \tag{16}$$

$$\mathsf{C}_2(\mathcal{S}) = \mathsf{C}\left(\sum_{k\in\mathcal{S}} \frac{|G_k|^2 P_r}{N}\right), \tag{17}$$

$$\mathsf{C}_3(\mathcal{S}) = \sum_{k\in\mathcal{S}} \mathsf{C}\left(\frac{1}{Q_k}\right), \tag{18}$$

the maximization is taken over all $\alpha_k \in (0, 1]$, $Q_k > 0$, $k \in [1 : N]$, and the distribution of $|\tilde{H}_k|^2$ is given in (4). Note that each term in (15) can be simplified as

$$\begin{aligned}
&\sum_{\Lambda\subseteq[1:N]} \left(\prod_{j\in[1:N]\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda)\right] \\
&\quad\overset{(a)}{=} \sum_{\Lambda\subseteq\mathcal{S}^c} \left(\prod_{j\in\mathcal{S}^c\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda)\right], \\
&\sum_{\Lambda\subseteq[1:N]} \left(\prod_{j\in[1:N]\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_2(\mathcal{S})\right] \\
&\quad= \mathsf{E}\left[\mathsf{C}_2(\mathcal{S})\right], \\
&\sum_{\Lambda\subseteq[1:N]} \left(\prod_{j\in[1:N]\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_3(\mathcal{S}\setminus\Lambda)\right] \\
&\quad= \sum_{\Lambda\subseteq\mathcal{S}} \left(\prod_{j\in\mathcal{S}\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{C}_3(\mathcal{S}\setminus\Lambda),
\end{aligned} \tag{19}$$

where step (a) follows from

$$\begin{aligned}
&\sum_{\Lambda\subseteq[1:N]} \left(\prod_{j\in[1:N]\setminus\Lambda} \alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda)\right] \\
&= \sum_{\Lambda_1\subseteq\mathcal{S}^c,\Lambda_2\subseteq\mathcal{S}} \left(\prod_{j\in\mathcal{S}^c\setminus\Lambda_1} \alpha_j\right)\left(\prod_{j\in\mathcal{S}\setminus\Lambda_2} \alpha_j\right) \\
&\quad\times\left(\prod_{j\in\Lambda_1}(1-\alpha_j)\right)\left(\prod_{j\in\Lambda_2}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda_1)\right] \\
&= \left(\sum_{\Lambda_1\subseteq\mathcal{S}^c} \left(\prod_{j\in\mathcal{S}^c\setminus\Lambda_1} \alpha_j\right)\left(\prod_{j\in\Lambda_1}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda_1)\right]\right) \\
&\quad\times\left(\sum_{\Lambda_2\subseteq\mathcal{S}} \left(\prod_{j\in\mathcal{S}\setminus\Lambda_2} \alpha_j\right)\left(\prod_{j\in\Lambda_2}(1-\alpha_j)\right)\right) \\
&= \sum_{\Lambda_1\subseteq\mathcal{S}^c} \left(\prod_{j\in\mathcal{S}^c\setminus\Lambda_1} \alpha_j\right)\left(\prod_{j\in\Lambda_1}(1-\alpha_j)\right)\mathsf{E}\left[\mathsf{C}_1(\mathcal{S}^c\setminus\Lambda_1)\right],
\end{aligned}$$

and by rewriting Λ_1 with Λ. In the same manner, we can prove the last two equalities in (19). Thus,

$$C \geq \max \min_{\mathcal{S}\subseteq[1:N]} R_{\text{ONNC}}(\mathcal{S}) \tag{20}$$

where $R_{\text{ONNC}}(\mathcal{S})$ is defined in (6).

Before presenting the proof of the asymptotic capacity result using the adaptive noisy network coding scheme, for comparison, we first present how the (non-adaptive) noisy network coding performs for fading networks. By fixing the compression noise variance Q_k independent of its first-hop channel gains, i.e., $\eta_k(h_k) = Q_k$, $k \in [1:N]$, the achievable rate of the non-adaptive noisy network coding scheme is given by

$$C \geq \min_{\mathcal{S}\subseteq[1:N]} \left\{ \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c} \frac{|H_k|^2 P}{1+Q_k} \right) \right] + \mathsf{C}_2(\mathcal{S}) - \mathsf{C}_3(\mathcal{S}) \right\}. \quad (21)$$

To see an example on the performance of this scheme, let $Q_k = 1$. Then, (21) becomes

$$C \geq \min_{\mathcal{S}\subseteq[1:N]} \left\{ \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c} |H_k|^2 \frac{P}{2} \right) - \mathsf{C}_2(\mathcal{S}) \right] - |\mathcal{S}| \right\} \quad (22)$$

where $|\mathcal{S}|$ denotes the cardinality of $\mathcal{S}$. By comparing (22) with the cut-set upper bound (7) it can be shown that noisy network coding scheme achieves within N bits/s/Hz of the capacity, independent of $\sigma_{H_k}, \sigma_{G_k}, P$, and P_r. The above result extends the capacity gap result of [5] for Gaussian (non-fading) networks to fading parallel relay networks. The result of [5] is a general purpose bound (for a general topology) of $1.26N$ bits/s/Hz. Here, we improve this bound to N bits/s/Hz by taking advantage of the specific topology and the fact that the cut-set upper bound is maximized over a product distribution for fading parallel networks. This type of performance guarantee is appealing in the high SNR regime. However, the capacity gap result does not say much when the number of relays is large due to the unbounded capacity gap in the limit of large N.

B. Asymptotic Capacity for Symmetric Fading

In this subsection, we prove Theorem 2 by showing that the proposed threshold-based opportunistic noisy network coding scheme can achieve the capacity in the symmetric setting as the number of relays becomes large. This result demonstrates how the opportunistic gain can improve the overall network performance by utilizing the channel state information at the receiver side.

Consider the symmetric case where $\sigma_{H_k} = \sigma_{G_k} = 1$. In the following, we first prove $\lim_{N\to\infty} C_N \leq \mathsf{C}(P_r)$ and then $\lim_{N\to\infty} C_N \geq \mathsf{C}(P_r)$.

From (7), the cut-set upper bound can be further bounded by

$$\begin{aligned} C_N &\leq \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{N} \frac{|G_k|^2 P_r}{N} \right) \right] \\ &\leq \mathsf{C}(P_r) \end{aligned} \quad (23)$$

where the second inequality holds from Jensen's inequality. Since (23) holds for any N, we have $\lim_{N\to\infty} C_N \leq \mathsf{C}(P_r)$.

Now consider the achievable rate of the threshold-based scheme in Theorem 1. By symmetry, we set $\alpha_k = \alpha$ (equivalently, $\gamma_k = \gamma$) and $Q_k = Q$ for all $k \in [1:N]$. Then the original 2^N rate constraints in Theorem 1 simplify to $N+1$ rate constraints by noticing the fact that the rate constraints corresponding to cuts with the same cardinality are all equivalent. Then, after some manipulation, we can show that

$$C_N \geq \max_{\alpha\in(0,1],Q>0} \min_{i\in[0:N]} R_{\mathrm{ONNC}}(i) \quad (24)$$

where

$$\begin{aligned} R_{\mathrm{ONNC}}(i) = &\sum_{j=0}^{i-1} \binom{i}{j} \alpha^{i-j}(1-\alpha)^j \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{i-j} \frac{|\tilde{H}_k|^2 P}{1+Q} \right) \right] \\ &+ \mathsf{E}\left[\mathsf{C}\left(\sum_{k=i+1}^{N} \frac{|G_k|^2 P_r}{N} \right) \right] - \alpha(N-i)\mathsf{C}\left(\frac{1}{Q}\right). \end{aligned} \quad (25)$$

In the following we show that $\lim_{N\to\infty} C_N \geq \mathsf{C}(P_r)$. Let $\alpha = \frac{\log\log(N)}{N}$ and $Q = \frac{P}{P_r}\ln(N)$. Then γ is given by $\ln(N/\log\log(N))$. First, consider the case where $i \in [\lceil N/\sqrt{\log\log(N)}\rceil, N]$. For this case, we bound $R_{\mathrm{ONNC}}(i)$ as follows:

$$\begin{aligned} &R_{\mathrm{ONNC}}(i) \\ &\overset{(a)}{\geq} \sum_{j=0}^{i-1} \binom{i}{j} \alpha^{i-j}(1-\alpha)^j \mathsf{C}\left(\frac{\gamma P}{1+Q}\right) - \alpha N \mathsf{C}\left(\frac{1}{Q}\right) \\ &\overset{(b)}{=} \left(1-(1-\alpha)^i\right) \mathsf{C}\left(\frac{\gamma P}{1+Q}\right) - \alpha N \mathsf{C}\left(\frac{1}{Q}\right) \\ &\overset{(c)}{\geq} \left(1 - \left(1 - \frac{\log\log(N)}{N}\right)^{\left\lceil \frac{N}{\sqrt{\log\log(N)}} \right\rceil} \right) \\ &\quad \times \mathsf{C}\left(\frac{\ln\left(N/\log\log(N)\right) P}{1+\frac{P}{P_r}\ln(N)} \right) - \log(e)\frac{P_r}{P}\frac{\log\log(N)}{\ln(N)} \\ &\overset{(d)}{\geq} \left(1 - \left(1 - \frac{\log\log(N)}{N}\right)^{\frac{N}{\log\log(N)} \frac{\log\log(N)}{\sqrt{\log\log(N)}}} \right) \\ &\quad \times \mathsf{C}\left(\frac{\ln\left(N/\log\log(N)\right) P}{1+\frac{P}{P_r}\ln(N)} \right) - \log(e)\frac{P_r}{P}\frac{\log\log(N)}{\ln(N)} \end{aligned} \quad (26)$$

where (a) follows since

$$\begin{aligned} \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{i-j} \frac{|\tilde{H}_k|^2 P}{1+Q} \right) \right] &\geq \mathsf{E}\left[\mathsf{C}\left(\frac{|\tilde{H}_1|^2 P}{1+Q} \right) \right] \\ &\geq \mathsf{C}\left(\frac{\gamma P}{1+Q}\right), \end{aligned}$$

for $j \in [0:i-1]$, step (b) follows from the fact that $\sum_{j=0}^{i}\binom{i}{j}\alpha^{i-j}(1-\alpha)^j = 1$, step (c) follows since $i = \lceil N/\sqrt{\log\log(N)}\rceil$ gives the minimum value and $\mathsf{C}(x) \leq x\log(e)$, and step (d) follows from $\lceil N/\sqrt{\log\log(N)}\rceil \geq N/\sqrt{\log\log(N)}$. Next, consider the case where $i \in [0 : \lfloor N/\sqrt{\log\log(N)}\rfloor]$. Similarly, we have

$$\begin{aligned} R_{\mathrm{ONNC}}(i) &\geq \mathsf{E}\left[\mathsf{C}\left(\sum_{k=i+1}^{N} \frac{|G_k|^2 P_r}{N} \right) \right] - \alpha N \mathsf{C}\left(\frac{1}{Q}\right) \\ &\geq \mathsf{E}\left[\mathsf{C}\left(\sum_{k=\left\lfloor N/\sqrt{\log\log(N)}\right\rfloor+1}^{N} \frac{|G_k|^2 P_r}{N} \right) \right] \\ &\quad - \log(e)\frac{P_r}{P}\frac{\log\log(N)}{\ln(N)}. \end{aligned} \quad (27)$$

Hence, from (24) to (27),

$$C_N \geq \min\left\{\left(1-\left(1-\frac{\log\log(N)}{N}\right)^{\frac{N}{\log\log(N)}\frac{\log\log(N)}{\sqrt{\log\log(N)}}}\right)\right. \times \mathsf{C}\left(\frac{\ln(N/\log\log(N))\,P}{1+\frac{P}{P_r}\ln(N)}\right), \left.\mathsf{E}\left[\mathsf{C}\left(\sum_{k=\left\lfloor N/\sqrt{\log\log(N)}\right\rfloor+1}^{N}\frac{|G_k|^2P_r}{N}\right)\right]\right\} - \log(e)\frac{P_r}{P}\frac{\log\log(N)}{\ln(N)}. \tag{28}$$

Finally, from

$$\lim_{N\to\infty} 1-\left(1-\frac{\log\log(N)}{N}\right)^{\frac{N}{\log\log(N)}\frac{\log\log(N)}{\sqrt{\log\log(N)}}} = 1, \tag{29}$$

$$\lim_{N\to\infty}\frac{\ln(N/\log\log(N))\,P}{1+\frac{P}{P_r}\ln(N)} = P_r,$$

$$\lim_{N\to\infty}\frac{\log\log(N)}{\ln(N)} = 0,$$

$$\lim_{N\to\infty}\frac{1}{N-\left\lfloor\frac{N}{\sqrt{\log\log(N)}}\right\rfloor}\sum_{k=\left\lfloor\frac{N}{\sqrt{\log\log(N)}}\right\rfloor+1}^{N}|G_k|^2 = 1 \tag{30}$$

we have $\lim_{N\to\infty} C_N \geq \mathsf{C}(P_r)$. Here, $\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^x = \frac{1}{e}$ is used in (29), and (30) holds from the law of large numbers. In conclusion, Theorem 2 holds.

Remark 2: As shown in the proof above, the choice $Q = \frac{P}{P_r}\ln(N)$ and $\gamma = \ln(N/\log\log(N))$ is the asymptotically optimal choice for the opportunistic noisy network coding scheme. The interpretation of this choice is that the threshold value should be increased, while each relay compression is set to be coarse as the number of relays increases.

V. Comparison

In this section, we compare the opportunistic noisy network coding scheme with AF and DF. For the AF scheme, it may not be always beneficial to forward every received symbol at the relays, but instead it may be better to forward a subset of received symbols while boosting the relay power. To be fair, we adopt a similar opportunistic concept to the AF scheme in which each relay only amplify-forwards a subset of received symbols with channel gains above a certain threshold. On the other hand, since the whole source message is recovered at each relay in the DF scheme, it is structurally impossible for opportunistic transmission based on CSIR using a similar concept. However, since the decode-forward scheme can always choose to use only a subset of the relays nodes, we compare with the decode-forward scheme that uses the optimal subset of relays.

A. Amplify-Forward Relaying

As mentioned above, a similar adaptation used in Section IV can also be applied to AF relaying. Specifically, relay node $k \in [1:N]$ sends $X_k = \zeta_k(H_k)Y_k$, where $\zeta_k(h_k)$ is given by

$$\zeta_k(h_k) = \begin{cases}\sqrt{\frac{P_r/(\alpha_k N)}{|h_k|^2P+1}} & \text{if } |h_k|^2 \geq \gamma_k,\\ 0 & \text{otherwise,}\end{cases} \tag{31}$$

which satisfies the power constraints. Then the opportunistic AF scheme results in the following lower bound:

$$\begin{aligned} C &\geq \max \mathsf{E}\left[\mathsf{C}\left(\frac{\left|\sum_{k=1}^N G_kH_k\zeta_k(H_k)\right|^2 P}{\sum_{k=1}^N|G_k|^2\zeta_k^2(H_k)+1}\right)\right]\\ &= \max\sum_{\Lambda\subseteq[1:N]}\left(\prod_{j\in[1:N]\setminus\Lambda}\alpha_j\right)\left(\prod_{j\in\Lambda}(1-\alpha_j)\right)\\ &\quad\times\mathsf{E}\left[\mathsf{C}\left(\frac{\left|\sum_{k\in[1:N]\setminus\Lambda}G_k\tilde{H}_k\sqrt{\frac{P_r/(\alpha_kN)}{|\tilde{H}_k|^2P+1}}\right|^2P}{\sum_{k\in[1:N]\setminus\Lambda}\frac{|G_k|^2P_r/(\alpha_kN)}{|\tilde{H}_k|^2P+1}+1}\right)\right] \end{aligned} \tag{32}$$

where the maximization is taken over all $\alpha_k \in (0,1]$, $k \in [1:N]$. This opportunistic AF scheme generalizes the conventional AF scheme, which corresponds to the case when $\alpha_k = 1$ for all $k \in [1:N]$. Similarly, for the symmetric case in which $\sigma_{H_k} = 1$ and $\sigma_{G_k} = 1$, we have

$$\begin{aligned} C &\geq \max_{\alpha\in(0,1]}\sum_{j=0}^{N-1}\binom{N}{j}\alpha^{N-j}(1-\alpha)^j\\ &\quad\times\mathsf{E}\left[\mathsf{C}\left(\frac{\left|\sum_{k=1}^{N-j}G_k\tilde{H}_k\sqrt{\frac{P_r/(\alpha N)}{|\tilde{H}_k|^2P+1}}\right|^2P}{\sum_{k=1}^{N-j}\frac{|G_k|^2P_r/(\alpha N)}{|\tilde{H}_k|^2P+1}+1}\right)\right]. \end{aligned} \tag{33}$$

The following theorem shows an upper bound on the achievable rate of the opportunistic AF scheme for the symmetric case.

Theorem 3: Consider the fading parallel relay network with $\sigma_{H_k} = \sigma_{G_k} = 1$. Then the achievable rate of the opportunistic AF scheme is upper bounded by $\mathsf{E}[\mathsf{C}(|g_1|^2P_r)]$ for any P and P_r.

Proof: Denote the right hand side of (33) as R_{AF}. Then

$$\begin{aligned} R_{AF} &\leq \max_{\alpha\in(0,1]}\sum_{j=0}^{N-1}\binom{N}{j}\alpha^{N-j}(1-\alpha)^j\\ &\quad\times\mathsf{E}\left[\mathsf{C}\left(\left|\sum_{k=1}^{N-j}G_k\tilde{H}_k\sqrt{\frac{P_r/(\alpha N)}{|\tilde{H}_k|^2P}}\right|^2P\right)\right]\\ &= \max_{\alpha\in(0,1]}\sum_{j=0}^{N-1}\binom{N}{j}\alpha^{N-j}(1-\alpha)^j\\ &\quad\times\mathsf{E}\left[\mathsf{C}\left(\left|\frac{1}{\sqrt{N-j}}\sum_{k=1}^{N-j}G_k\frac{\tilde{H}_k}{|\tilde{H}_k|}\right|^2\frac{(N-j)P_r}{\alpha N}\right)\right]\\ &= \max_{\alpha\in(0,1]}\sum_{j=0}^{N}\binom{N}{j}\alpha^{N-j}(1-\alpha)^j\\ &\quad\times\mathsf{E}\left[\mathsf{C}\left(|G_1|^2\frac{(N-j)P_r}{\alpha N}\right)\right] \end{aligned} \tag{34}$$

where the last equality holds since the probability distribution of $G_k \frac{\tilde{H}_k}{|\tilde{H}_k|}$ is given by $\mathcal{N}_{\mathbb{C}}(0, 1)$. Note that $\sum_{j=0}^{N} \binom{N}{j} \alpha^{N-j} (1-\alpha)^j = 1$ and $\sum_{j=0}^{N} \binom{N}{j} \alpha^{N-j} (1-\alpha)^j \frac{(N-j)P_r}{\alpha N} = P_r$ for any $\alpha \in (0, 1]$. Hence R_{AF} is upper bounded by the following:

$$R_{AF} \leq \max_{\substack{\{q_j \geq 0\}_{j=0}^N, \{Q_j \geq 0\}_{j=0}^N, \\ \sum_{j=0}^N q_j = 1, \sum_{j=0}^N q_j Q_j = P_r}} \sum_{j=0}^{N} q_j \mathsf{E}\left[\mathsf{C}\left(|G_1|^2 Q_j\right)\right]. \quad (35)$$

Since $\mathsf{E}[\mathsf{C}(|G_1|^2 x)]$ is a concave function on $x \geq 0$, $Q_j = P_r$ for all $j \in [0 : N]$ maximizes (35), which gives $R_{AF} \leq \mathsf{E}[\mathsf{C}(|G_1|^2 P_r)]$. □

Unlike the opportunistic noisy network coding scheme, Theorem 3 states that the opportunistic AF scheme cannot achieve the capacity of fading symmetric parallel relay networks even if $N \to \infty$ since $\mathsf{E}[\mathsf{C}(|G_1|^2 P_r)]$ is strictly less than $\mathsf{C}(P_r)$.

B. Decode-Forward Relaying

For the decode-forward strategy, suppose that only the relays in $\mathcal{S} \subseteq [1 : N]$ decode the message and participate in the second-hop transmission. Due to the decoding constraints at the relays, the rate of the DF scheme is limited by the minimum of the point-to-point capacities between the source and each of the relays in $\mathcal{S}$, which gives

$$C \geq \max_{\mathcal{S} \subseteq [1:N]} \min \left\{ \min_{k \in \mathcal{S}} \mathsf{E}\left[\mathsf{C}\left(|H_k|^2 P\right)\right], \mathsf{E}\left[\mathsf{C}_2(\mathcal{S})\right] \right\}, \quad (36)$$

where $\mathsf{C}_2(\mathcal{S})$ is defined in (17). For the symmetric case in which $\sigma_{H_k} = 1$ and $\sigma_{G_k} = 1$, the above rate is simplified to

$$C \geq \min \left\{ \mathsf{E}\left[\mathsf{C}\left(|H_1|^2 P\right)\right], \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{N} |G_k|^2 \frac{P_r}{N}\right)\right] \right\}. \quad (37)$$

Hence, from the cut-set upper bound (7), DF achieves the capacity if

$$\mathsf{E}\left[\mathsf{C}\left(|H_1|^2 P\right)\right] \geq \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{N} |G_k|^2 \frac{P_r}{N}\right)\right] \quad (38)$$

for the symmetric case. For high SNR and the large number of relays regime, the optimality condition (38) is approximately given by

$$P \geq 2^{0.83} P_r \quad (39)$$

where $\mathsf{E}\left[\mathsf{C}(|H_1|^2 P)\right] \simeq \log(P) - 0.83$ and $\lim_{N\to\infty} \mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{N} |G_k|^2 \frac{P_r}{N}\right)\right] \simeq \log(P_r)$ are used. On the other hand, if $P \leq P_r$, the right hand side of (37) is given by $\mathsf{E}[\mathsf{C}(|H_1|^2 P)]$, which is strictly less than $\mathsf{C}(P_r)$. Hence it cannot achieve the capacity for this case.

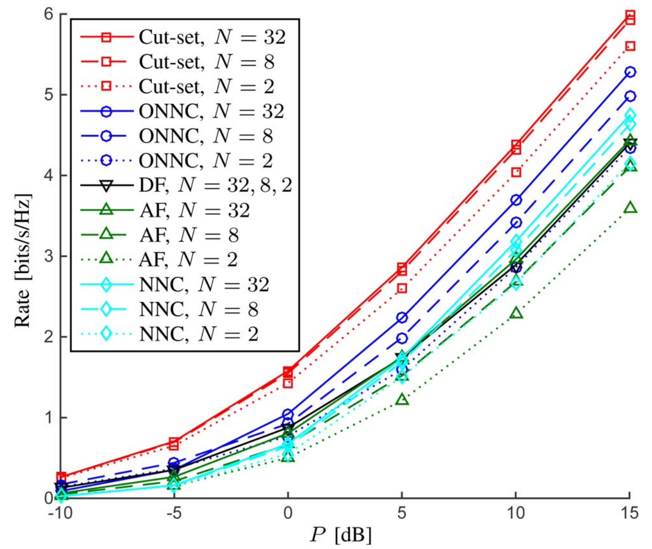

Fig. 3. Achievable rates for the symmetric case when $P_r = 2P$ for $N = 2, 8, 32$.

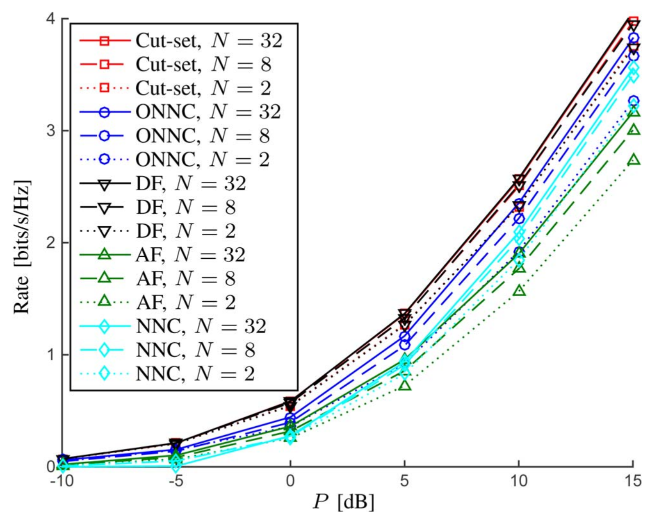

Fig. 4. Achievable rates for the symmetric case when $P_r = 0.5P$ for $N = 2, 8, 32$.

C. Rate Comparison

1) Symmetric Networks: In this subsection, we first compare the achievable rate of the opportunistic noisy network coding scheme with those of the AF and DF schemes for the symmetric case by numerical evaluation of (24), (32), and (37), respectively. We also compare with the non-adaptive noisy network coding in (21) by setting $Q_k = Q$ and optimizing with Q.

Fig. 3 plots the achievable rates when $P_r = 2P$. As shown in the figure, opportunistic noisy network coding outperforms the other schemes in most cases, and the rate gap from the cut-set upper bound decreases as the number of relays increases. However, opportunistic noisy network coding does not always outperform DF. As intuition suggests, for the case where the SNR of the first hop is higher than the second hop, DF can be better than the opportunistic noisy network coding scheme. Fig. 4 plots the achievable rates when $P_r = 0.5P$. Since this

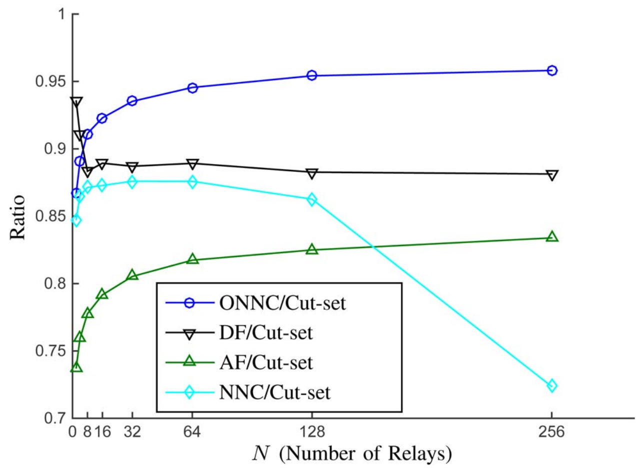

Fig. 5. Ratios between achievable rates and the cut-set upper bound for the symmetric case when $P = P_r = 20$ dB.

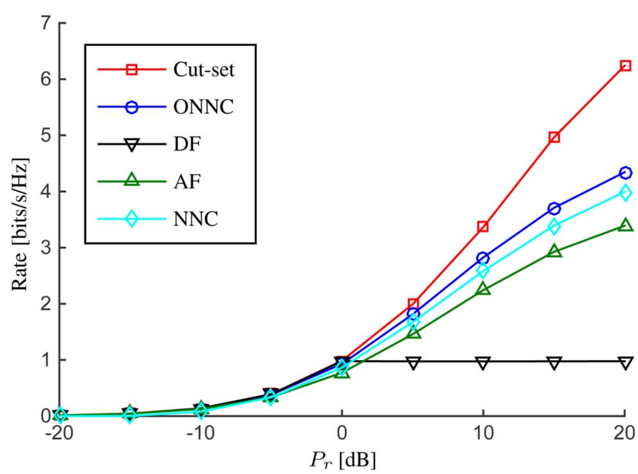

Fig. 6. Achievable rates for the symmetric case when $P = 10$ dB and $N = 8$.

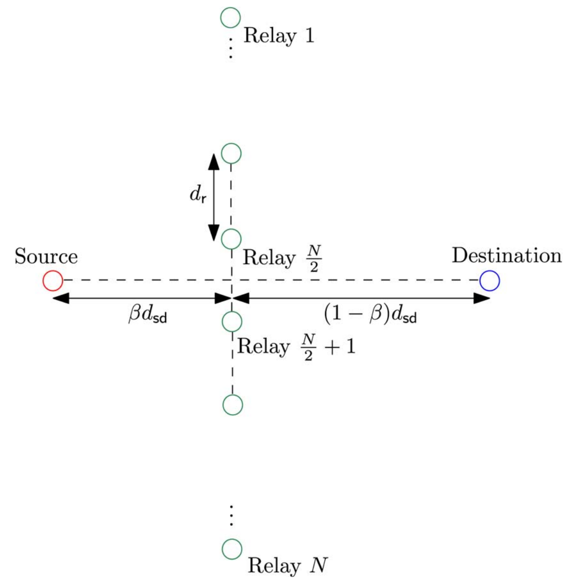

Fig. 7. Geometric network model consisting of a set of relays regularly deployed on a line.

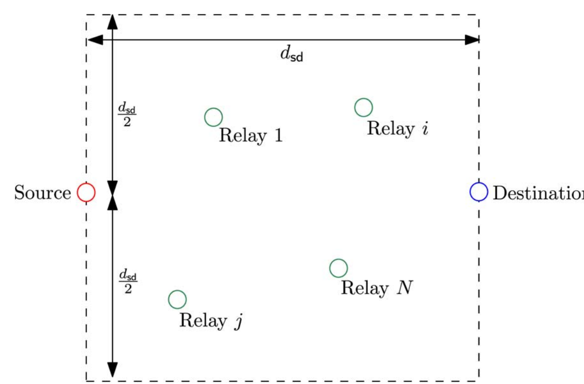

Fig. 8. Geometric network model consisting of a set of relays randomely deployed in a square area.

regime is close to the optimality condition of DF in (38), the rate of DF is very close to the cut-set upper bound. Although DF is better for such cases, opportunistic noisy network coding eventually converges to the cut-set upper bound, as verified in Theorem 2.

Fig. 5 plots the ratios between the achievable rates and the cut-set upper bound with respect to the number of relays when $P = P_r = 20$ dB. Although the convergence of the opportunistic noisy network coding rate and the cut-set bound requires the use of many relays, the gap from the cut-set upper bound decreases and eventually converges to zero as N increases. On the other hand, as shown in Sections V-A and V-B, AF and DF cannot achieve the capacity even if $N \to \infty$. Perhaps more important than this unique convergence property over the alternative schemes compared here, we can see that opportunistic noisy network coding provides significant gain over the other schemes, especially over the non-adaptive noisy network coding scheme.

Lastly, as shown in Figs. 4 and 5, the achievable rate of each scheme is affected by the ratio between P and P_r. Fig. 6 plots the achievable rates with respect to P_r when $P = 10$ dB. For a wide SNR range of interest, opportunistic noisy network coding provides an improved rate compared to the other schemes.

2) Asymmetric Networks: In order to verify the rate gain of opportunistic noisy network coding for general asymmetric networks in a meaningful manner, we consider two geometric network configurations depicted in Figs. 7 and 8.

For the first model in Fig. 7, a set of N relays are regularly deployed on a line with distance d_{r} to each other. For convenience, we assume N is even for the first model. The distance between the source and destination is given by d_{sd} and the relay line is located at a distance of βd_{sd} from the source. Hence the distance between the source and relay $k \in [1:N]$ is given by $d_{\mathsf{sr},k} = \sqrt{\beta^2 {d_{\mathsf{sd}}}^2 + \left(k - \frac{N+1}{2}\right)^2 {d_{\mathsf{r}}}^2}$ and the distance between relay k and the destination is given by $d_{\mathsf{rd},k} = \sqrt{(1-\beta)^2 {d_{\mathsf{sd}}}^2 + \left(k - \frac{N+1}{2}\right)^2 {d_{\mathsf{r}}}^2}$. The path-loss channel model is assumed in which the average received signal power decreases as $d^{-\mu}$ when the transmit distance is given by d, where $\mu \geq 2$ is the path-loss exponent. Therefore, the

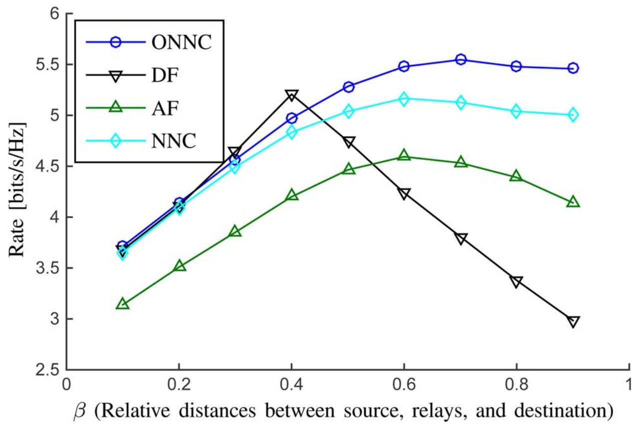

Fig. 9. Achievable rates for the geometric network in Fig. 7 when $P = P_r = 10$ dB and $N = 8$.

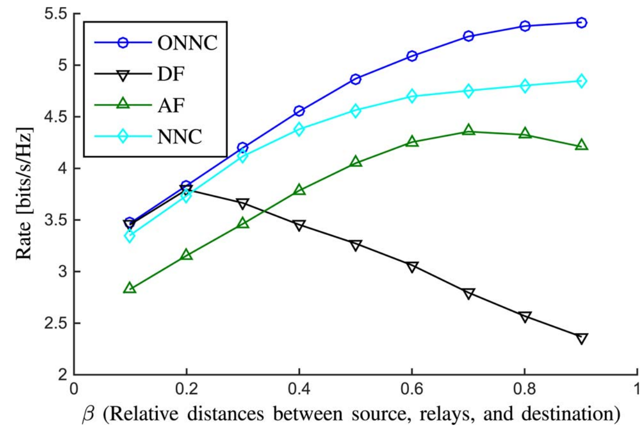

Fig. 10. Achievable rates for the geometric network in Fig. 7 when $P = P_r = 10$ dB and $N = 16$.

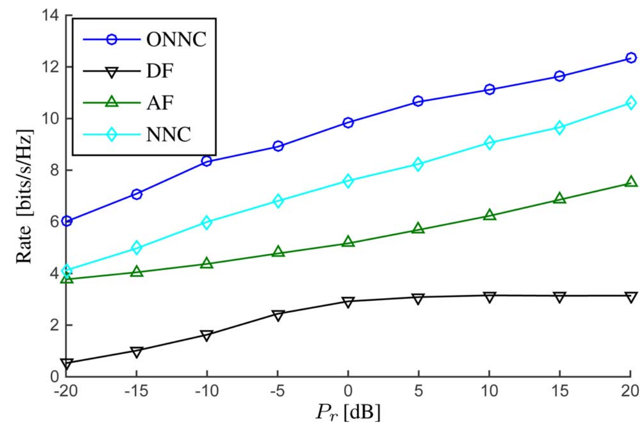

Fig. 11. Achievable rates for the geometric network in Fig. 8 when $P = 10$ dB and $N = 8$.

channel coefficients from the source to relay k and from relay k to the destination at time t are given by

$$H_{\mathsf{geo},k}[t] = \frac{H_k[t]}{(d_{\mathsf{sr},k})^{\mu/2}} \tag{40}$$

and

$$G_{\mathsf{geo},k}[t] = \frac{G_k[t]}{(d_{\mathsf{rd},k})^{\mu/2}}, \tag{41}$$

respectively. Here $H_k[t]$ and $G_k[t]$ are fading components defined in Section II-A with $\sigma^2_{H_k} = \sigma^2_{G_k} = 1$.

For the second model in Fig. 8, a set of N relays are uniformly deployed at random in a $d_{\mathsf{sd}} \times d_{\mathsf{sd}}$ square area. Suppose that the source and destination are located at $\left(0, \frac{d_{\mathsf{sd}}}{2}\right)$ and $\left(d_{\mathsf{sd}}, \frac{d_{\mathsf{sd}}}{2}\right)$ respectively and the location of relay k is given by $(x, y) \in [0, d_{\mathsf{sd}}]^2$. Then the channel coefficients from the source to relay k and from relay k to the destination at time t are given by (40) and (41) respectively, where $d_{\mathsf{sr},k} = \sqrt{x^2 + \left(\frac{d_{\mathsf{sd}}}{2} - y\right)^2}$ and $d_{\mathsf{rd},k} = \sqrt{(x - d_{\mathsf{sd}})^2 + \left(\frac{d_{\mathsf{sd}}}{2} - y\right)^2}$.

We compare the achievable rates of opportunistic noisy network coding, AF, DF, and non-adaptive noisy network coding, given by (5), (32), (36), and (21) respectively, under the above two geometric models. For opportunistic noisy network coding, we set $\alpha_k = \alpha$, or equivalently $\gamma_k = \frac{\ln(1/\alpha)}{(d_{\mathsf{sr},k})^\mu}$, and $Q_k = Q$. Then we numerically optimize (5) for the simulations. In a similar manner, we optimize (32) with $\alpha_k = \alpha$ and optimize (21) with $Q_k = Q$. In these numerical simulations, we assume $d_{\mathsf{sd}} = 1$, $d_{\mathsf{r}} = 0.1$, and $\mu = 3$, but similar rate performance can be observed for various different network configurations.

Figs. 9 and 10 plot the achievable rates for the first geometric network model with respect to β when $N = 8$ and $N = 16$, respectively. As β increases, i.e., relays are closer to the destination, opportunistic noisy network coding outperforms AF, DF, and non-adaptive noisy network coding. Fig. 11 plots the achievable rates for the second geometric network model with respect to P_r when $P = 10$ dB and $N = 8$. Similar to the symmetric case in Fig. 6, opportunistic noisy network coding outperforms the other schemes for a wide SNR range.

VI. GENERALIZATIONS

In this section, we briefly discuss some possible generalizations to other fading distributions and network configurations.

A. General Channel Distributions

Although we have focused on Rayleigh fading in the previous sections, the results presented in this paper can be extended to more general channel distributions. Obviously, the cut-set upper bound (7) and the opportunistic noisy network coding lower bound in Theorem 1 apply to any channel distributions. Furthermore, Theorem 3 also hold for any channel distributions since the results are not limited to a specific channel distribution.

As for Theorem 2 which relies on the Rayleigh fading assumption, we can extend the theorem to a more general class of channel distributions. As before, we assume $\alpha_k = \alpha$ (equivalently, $\gamma_k = \gamma$) and $Q_k = Q$. Define a class of probability distributions on H_k such that there exists an increasing sequence $f(N) > 0$ with $\lim_{N\to\infty} \frac{f(N)}{\log(N)} = 0$ that satisfies $\lim_{N\to\infty} \frac{\alpha N}{\gamma} = 0$, where $\alpha = \frac{f(N)}{N}$. Note that Rayleigh distribution is included in this class where $f(N) = \log\log(N)$. For this class of probability distributions, we can show that the same result presented in Theorem 2 apply by following similar steps as in the Rayleigh fading case.

B. Two-Way Communications

Consider the fading two-way parallel relay network in which two nodes exchange messages with the help of N relays. For the two-way parallel relay channel, the input-output relations at time t are given by

$$Y_k[t] = H_{ka}[t]X_a[t] + H_{kb}[t]X_b[t] + Z_k[t],$$
$$Y_a[t] = \sum_{k=1}^{N} G_{ak}[t]X_k[t] + Z_a[t],$$
$$Y_b[t] = \sum_{k=1}^{N} G_{bk}[t]X_k[t] + Z_b[t]. \tag{42}$$

We assume average power constraint P on both source nodes and P_r/N for each relay node. The channel coefficients of the links from the source nodes a and b to the relays are given by $H_{ka}[t]$ and $H_{kb}[t]$, respectively. Similarly, the the channel coefficients of the links from the relays to each source node is given by $G_{ak}[t]$ and $G_{bk}[t]$. We assume that all channel coefficients are independent zero mean Gaussian random variables as in the one-way relay network case. Note that this model is the fading version of the Gaussian two-way channel if $N = 2$, which has been extensively studied in the literature [12], [26], [27].

As in the one-way parallel relay case, we show that the opportunistic noisy network coding scheme achieves the capacity region of the fading two-way symmetric parallel relay network as N increases. Since the overall proof is similar to that of Theorem 2, we provide an outline of the proof here.

Let (R_a, R_b) denote an achievable rate pair and $\eta_k(h_{ka}, h_{kb}) > 0$ denote the adaptation function of relay node k, $k \in [1 : N]$. Note that in the two-way relay case, each relay node adapts to a pair of channel gains. Then, similar to (13), a rate pair (R_a, R_b) satisfying

$$R_a \le \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c}\frac{|H_{ka}|^2P}{1+\eta_k(H_{ka},H_{kb})}\right) + \mathsf{C}\left(\sum_{k\in\mathcal{S}}|G_{bk}|^2\frac{P_r}{N}\right) - \sum_{k\in\mathcal{S}}\mathsf{C}\left(\frac{1}{\eta_k(H_{ka},H_{kb})}\right)\right],$$
$$R_b \le \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c}\frac{|H_{kb}|^2P}{1+\eta_k(H_{ka},H_{kb})}\right) + \mathsf{C}\left(\sum_{k\in\mathcal{S}}|G_{ak}|^2\frac{P_r}{N}\right) - \sum_{k\in\mathcal{S}}\mathsf{C}\left(\frac{1}{\eta_k(H_{ka},H_{kb})}\right)\right] \tag{43}$$

is achievable for some $\eta_k(h_{ka}, h_{kb}) > 0$, $k \in [1 : N]$. Consider the symmetric case in which the variances of all channel coefficients are equal to one. Let $\alpha = \mathsf{P}\{|H_{ka}|^2 \ge \gamma, |H_{kb}|^2 \ge \gamma\}$, which gives $\gamma = \frac{1}{2}\ln(1/\alpha)$. By setting

$$\eta_k(h_{ka}, h_{kb}) = \begin{cases} Q & \text{if } |h_{ka}|^2 \ge \gamma \text{ and } |h_{kb}|^2 \ge \gamma, \\ \infty & \text{otherwise} \end{cases} \tag{44}$$

with $Q > 0$, a rate pair (R_a, R_b) satisfying

$$R_a \le \min_{i\in[0:N]}\left\{\sum_{j=0}^{i-1}\binom{i}{j}\alpha^{i-j}(1-\alpha)^j\mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{i-j}\frac{|\tilde{H}_{ka}|^2P}{1+Q}\right)\right] + \mathsf{E}\left[\mathsf{C}\left(\sum_{k=i+1}^{N}\frac{|G_{bk}|^2P_r}{N}\right)\right] - \alpha(N-i)\mathsf{C}\left(\frac{1}{Q}\right)\right\},$$
$$R_b \le \min_{i\in[0:N]}\left\{\sum_{j=0}^{i-1}\binom{i}{j}\alpha^{i-j}(1-\alpha)^j\mathsf{E}\left[\mathsf{C}\left(\sum_{k=1}^{i-j}\frac{|\tilde{H}_{kb}|^2P}{1+Q}\right)\right] + \mathsf{E}\left[\mathsf{C}\left(\sum_{k=i+1}^{N}\frac{|G_{ak}|^2P_r}{N}\right)\right] - \alpha(N-i)\mathsf{C}\left(\frac{1}{Q}\right)\right\} \tag{45}$$

is achievable for some $\alpha \in (0, 1]$ and $Q > 0$. Here, $\tilde{H}_{ka}$ and $\tilde{H}_{kb}$ is defined similar to the definition of $\tilde{H}_k$ in the one-way case. Then by following the steps of the proof of Theorem 2 with $\alpha = \frac{\log\log(N)}{N}$ and $Q = \frac{P}{2P_r}\ln(N)$, it can be shown that $\lim_{N\to\infty} R_a = \mathsf{C}(P_r)$ and $\lim_{N\to\infty} R_b = \mathsf{C}(P_r)$. Therefore, in the limit of large N, the capacity region of the fading two-way symmetric parallel relay network is given by all rate pairs (R_a, R_b) such that

$$R_a < \mathsf{C}(P_r), \tag{46}$$
$$R_b < \mathsf{C}(P_r), \tag{47}$$

which is achievable by opportunistic noisy network coding.

C. Multicast Networks

Our model can further be generalized to the multicast network in which the source wishes to send a multicast message to the set of K destinations with the help of N relays. As in the single destination case, the channel from the source to the relays is a broadcast channel while the channel from the relays to each destination is a multiple access channel. Let H_1 to H_N be the channels from the source to the set of relay nodes and G_{k1} to G_{kN} be the channel gains from the relay nodes to destination k, $k \in [1 : K]$. Then by the union of events bound, the lower bound (13) can be generalized by

$$C \ge \max \min_{k\in[1:K]} \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\left[\mathsf{C}\left(\sum_{l\in\mathcal{S}^c}\frac{|H_l|^2P}{1+\eta_l(H_l)}\right) + \mathsf{C}\left(\sum_{l\in\mathcal{S}}|G_{kl}|^2\frac{P_r}{N}\right) - \sum_{l\in\mathcal{S}}\mathsf{C}\left(\frac{1}{\eta_l(H_l)}\right)\right] \tag{48}$$

where the maximization is taken over all $\eta_k(h_k) > 0$. Also, for the symmetric setting in which $\sigma_{H_j} = \sigma_H$ and $\sigma_{G_{kj}} = \sigma_{G_k}$ for all $j \in [1 : N]$ and $k \in [1 : N]$ we can show that by using the same threshold adaptation function in (14),

$$\lim_{N\to\infty} C_N = \min_{k\in[1:K]} \mathsf{C}\left(\sigma_k^2 P_r\right), \tag{49}$$

which implies that the opportunistic noisy network coding scheme achieves the capacity as $N \to \infty$ for the multicast case.

VII. Conclusion

In this paper, we proposed the opportunistic noisy network coding scheme for fading parallel relay networks in which each relay node selectively compresses the reliable received symbols that have channel gains above a certain threshold. Through this approach, the relays can efficiently transmit the reliable compressed symbols without wasting power on crude observations. In the symmetric case, the proposed scheme was shown to achieve the capacity in the limit of large number of relays. The optimal strategy is to compress fewer but better observations with higher channel gains as the number of relays increases. We further provided several detailed comparisons with competitive alternative schemes over geometric models and showed that our new scheme can strictly outperform both AF and DF for networks with large number of relays.

Appendix A
Probability of Error Analysis

In this section, we provide the probability of error analysis for the opportunistic noisy network coding scheme in Section IV-A. To deal with state (fading) dependent networks with state information at the receiver side, we use an argument similar to the one in [28]. Consider the augmented discrete memoryless network with channel outputs $\tilde{y}_k = (y_k, h_k)$ and $\tilde{y} = (y, g^N, h^N)$ where

$$p(\tilde{y}_1, \dots, \tilde{y}_N|x)p(\tilde{y}|x_1, \dots, x_N) = p(y^N|x, h^N)p(y|x_1, \dots, x_N, h^N, g^N)\prod_{k=1}^{N} p(h_k)p(g_k).$$

From the relation above, $\tilde{Y}_k$ and $\tilde{Y}$ are the channel outputs of the original channel. Note that the augmented network has $p(x, x^N, \tilde{y}^N, \tilde{y}) \neq p(x, \tilde{y}^N)p(x^N, \tilde{y})$, i.e., the layers in the augmented network are no longer independent. However, the layers are conditionally independent, i.e.,

$$p(x, x^N, \tilde{y}^N, \tilde{y}|h^N) = p(x, \tilde{y}^N|h^N)p(x^N, \tilde{y}|h^N). \tag{50}$$

We provide the probability of error analysis for recovering $m_j \in [1 : 2^{nR}]$ at the end of block $j+1$, $j \in [1 : b]$.

Let $M = M_j$ denote the message and $L_k = L_{kj}$, $k \in [1 : N]$, denote the index chosen by node k for block j. To bound the probability of error for the decoder, assume without loss of generality that $M = 1$ and $L_1 = \cdots = L_N = 1$. Then the decoder makes an error only if one of the following events occur:

$$\begin{aligned}
\mathcal{E}_1 &= \left\{\left(\hat{Y}^n_{kj}(l_k), \tilde{Y}^n_{kj}\right) \notin \mathcal{T}^{(n)}_{\epsilon'},\ \forall l_k \text{ for some } k \in [1 : N]\right\},\\
\mathcal{E}_{21} &= \left\{\left(X^n(1), \hat{Y}^n_{1j}(1), \dots, \hat{Y}^n_{Nj}(1), \mathbf{H}^N_j\right) \notin \mathcal{T}^{(n)}_{\epsilon}\right\},\\
\mathcal{E}_{22} &= \left\{\left(X^n_{1,j+1}(1), \dots, X^n_{N,j+1}(1), \tilde{Y}^n_{j+1}\right) \notin \mathcal{T}^{(n)}_{\epsilon}\right\},\\
\mathcal{E}_3 &= \Big\{\left(X^n_j(m), \hat{Y}^n_{1j}(l_1), \dots, \hat{Y}^n_{Nj}(l_N), \mathbf{H}^N_j\right) \in \mathcal{T}^{(n)}_{\epsilon},\\
&\qquad \left(X^n_{1,j+1}(l_1), \dots, X^n_{N,j+1}(l_N), \tilde{Y}^n_{j+1}\right) \in \mathcal{T}^{(n)}_{\epsilon},\\
&\qquad \text{for some } m \neq 1, l_1, \dots, l_N\Big\},
\end{aligned}$$

where $\mathbf{H}^N_j = (H^n_{1j}, \dots, H^n_{Nj})$. Thus, the probability of error is bounded as

$$\mathsf{P}(\mathcal{E}) \le \mathsf{P}(\mathcal{E}_1) + \mathsf{P}\left(\mathcal{E}_{21} \cap \mathcal{E}^c_1\right) + \mathsf{P}\left(\mathcal{E}_{22} \cap \mathcal{E}^c_1\right) + \mathsf{P}(\mathcal{E}_3).$$

By the covering lemma and the union of events bound, $\mathsf{P}(\mathcal{E}_1)$ tends to zero as $n \to \infty$ if

$$R_k > I(\hat{Y}_k; \tilde{Y}_k) + \delta(\epsilon'), \tag{51}$$

$k \in [1 : N]$. By the Markov lemma [25], the second term $\mathsf{P}(\mathcal{E}_{21} \cap \mathcal{E}^c_1)$ tends to zero as $n \to \infty$. By the conditional typicality lemma [25], the third term $\mathsf{P}(\mathcal{E}_{22} \cap \mathcal{E}^c_1)$ tends to zero as $n \to \infty$. For the final term, define the events

$$\begin{aligned}
\tilde{\mathcal{E}}_{1j}(m, l^N) &= \Big\{\Big(X^n_j(m), \hat{Y}^n_{1j}(l_1), \dots, \hat{Y}^n_{Nj}(l_N),\\
&\qquad H^n_{1j}, \dots, H^n_{Nj}\Big) \in \mathcal{T}^{(n)}_{\epsilon}\Big\},\\
\tilde{\mathcal{E}}_{2,j+1}(l^N) &= \left\{\left(X^n_{1,j+1}(l_1), \dots, X^n_{N,j+1}(l_N), \tilde{Y}^n_{j+1}\right) \in \mathcal{T}^{(n)}_{\epsilon}\right\}.
\end{aligned}$$

Then,

$$\begin{aligned}
\mathsf{P}(\mathcal{E}_3) &= \mathsf{P}\left(\bigcup_{m\neq 1, l^N} \tilde{\mathcal{E}}_{1j}(m, l^N) \cap \tilde{\mathcal{E}}_{2,j+1}(l^N)\right)\\
&\le \sum_{m\neq 1, l^N} \mathsf{P}\left(\tilde{\mathcal{E}}_{1j}(m, l^N), \tilde{\mathcal{E}}_{2,j+1}(l^N)\right)\\
&\le \sum_{m\neq 1, l^N} \mathsf{P}\left(\tilde{\mathcal{E}}_{1j}(m, l^N), \tilde{\mathcal{E}}_{2,j+1}(l^N), \mathbf{H}^N_j \in \mathcal{T}^{(n)}_{\epsilon}\right)\\
&\qquad + \mathsf{P}\left(\mathbf{H}^N_j \notin \mathcal{T}^{(n)}_{\epsilon}\right)\\
&\overset{(a)}{=} \sum_{m\neq 1, l^N} \sum_{\mathbf{h}^N \in \mathcal{T}^{(n)}_{\epsilon}} p(\mathbf{h}^N)\\
&\qquad \times \mathsf{P}\left(\tilde{\mathcal{E}}_{1j}(m, l^N), \tilde{\mathcal{E}}_{2,j+1}(l^N)|\mathbf{h}^N\right) + \epsilon_n\\
&\overset{(b)}{=} \sum_{m\neq 1, l^N} \sum_{\mathbf{h}^N \in \mathcal{T}^{(n)}_{\epsilon}} p(\mathbf{h}^N)\\
&\qquad \times \mathsf{P}\left(\tilde{\mathcal{E}}_{1j}(m, l^N)|\mathbf{h}^N\right) \mathsf{P}\left(\tilde{\mathcal{E}}_{2,j+1}(l^N)|\mathbf{h}^N\right) + \epsilon_n,
\end{aligned}$$

where $\mathbf{h}^N = (h^n_1, \dots, h^n_N)$, $\epsilon_n \to 0$ as $n \to \infty$, step (a) follows from the law of large numbers, and step (b) follows from (50). For l^N, let $\mathcal{S} = \mathcal{S}(l^N) = \{k \in [1 : N] : l_k \neq 1\}$ and $\mathcal{S}^c = \mathcal{S}^c(l^N) = \{k \in [1 : N] : l_k = 1\}$. Then, for $m \neq 1$, $(h^n_1, \dots, h^n_N) \in \mathcal{T}^{(n)}_{\epsilon}$, and some l^N index tuple,

$$\begin{aligned}
&\mathsf{P}\left(\tilde{\mathcal{E}}_{1j}(m, l^N)|\mathbf{h}^N\right)\\
&\quad = \mathsf{P}\left\{\left(X^n_j(m), \hat{Y}^n_{1j}(l_1), \dots, \hat{Y}^n_{Nj}(l_N), \mathbf{h}^N\right) \in \mathcal{T}^{(n)}_{\epsilon}|\mathbf{h}^N\right\}\\
&\quad = \sum_{(x^n, \hat{y}^n_1, \dots, \hat{y}^n_N) \in \mathcal{T}^{(n)}_{\epsilon}(X, \hat{Y}^N|\mathbf{h}^N)} p(x^n)p\left(\hat{y}^n(\mathcal{S})\right)p\left(\hat{y}^n(\mathcal{S}^c)|\mathbf{h}^N\right)\\
&\quad \le 2^{-n\left(I\left(X;\hat{Y}(\mathcal{S}^c)|H^N\right)+\sum_{k\in\mathcal{S}} I\left(\hat{Y}_k;\hat{Y}(\mathcal{S}_k),\hat{Y}(\mathcal{S}^c),X,H^N\right)-\delta(\epsilon)\right)},
\end{aligned}$$

where $\mathcal{S}_k = (\mathcal{S} \cap [1:k-1])$. On the other hand,

$$
\begin{aligned}
&\mathsf{P}\left(\tilde{\mathcal{E}}_{2j}(l^N)|\mathbf{h}^N\right)\\
&= \mathsf{P}\left\{\left(X^n_{1,j+1}(l_1), \ldots, X^n_{N,j+1}(l_N), \tilde{Y}^n_{j+1}\right) \in \mathcal{T}^{(n)}_\epsilon |\mathbf{h}^N\right\}\\
&= \sum_{(x^n_1,\ldots,x^n_N,\tilde{y}^n)\in\mathcal{T}^{(n)}_\epsilon(X^N,\tilde{Y}|\mathbf{h}^N)} p\left(\tilde{y}^n|x^n(\mathcal{S}^c), \mathbf{h}^N\right)\prod_{k=1}^N p\left(x^n_k\right)\\
&\le 2^{-n\left(I\left(X(\mathcal{S});\tilde{Y}|X(\mathcal{S}^c),H^N\right)-\delta(\epsilon)\right)}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\mathsf{P}(\mathcal{E}_3) &\le \sum_{m\neq 1}\sum_{l^N} 2^{-n(\mathsf{I}(\mathcal{S})-\delta(\epsilon))}\\
&\le \sum_{m\neq 1}\sum_{\mathcal{S}\subseteq[1:N]} 2^{n\hat{R}(\mathcal{S})}2^{-n(\mathsf{I}(\mathcal{S})-\delta(\epsilon))}\\
&\le \sum_{\mathcal{S}\subseteq[1:N]} 2^{n\left(R+\hat{R}(\mathcal{S})\right)}2^{-n(\mathsf{I}(\mathcal{S})-\delta(\epsilon))},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathsf{I}(\mathcal{S}) = I\left(X;\hat{Y}(\mathcal{S}^c)|H^N\right) &+ \sum_{k\in\mathcal{S}} I\left(\hat{Y}_k;\hat{Y}(\mathcal{S}_k),\hat{Y}(\mathcal{S}^c),X,H^N\right)\\
&+ I\left(X(\mathcal{S});\tilde{Y}|X(\mathcal{S}^c),H^N\right).
\end{aligned}
$$

Thus, the probability $\mathsf{P}(\mathcal{E}_3)$ tends to zero as $n\to\infty$ if

$$R < \min_{\mathcal{S}\subseteq[1:N]} \mathsf{I}(\mathcal{S}) - \hat{R}(\mathcal{S}) - \delta(\epsilon).$$

Finally, by rewriting $\tilde{Y} = (Y, G^N, H^N)$, using the fact that (Y, X^N) is independent of H^N, and the fact that the fact that X^N is independent of G^N, the above inequality simplifies to,

$$
\begin{aligned}
R < \min_{\mathcal{S}\subseteq[1:N]} & I\left(X;\hat{Y}(\mathcal{S}^c)|H^N\right)\\
&+ \sum_{k\in\mathcal{S}} I\left(\hat{Y}_k;\hat{Y}(\mathcal{S}_k),\hat{Y}(\mathcal{S}^c),X,H^N\right)\\
&+ I\left(X(\mathcal{S});Y|X(\mathcal{S}^c),G^N\right) - \hat{R}(\mathcal{S}) - \delta(\epsilon),
\end{aligned}
\tag{52}
$$

which concludes the proof.

APPENDIX B
CUT-SET UPPER BOUND

In this section, we prove the information-theoretic cut-set upper bound on the capacity of the fading parallel relay network. Similar to the cut-set upper bounds in [29], [30], the rate of any reliable length-n block coding should satisfy

$$R \le I\left(X, X(\mathcal{S}); Y(\mathcal{S}^c), Y|X(\mathcal{S}^c), H^N, G^N\right) + \epsilon_n \tag{53}$$

where $\epsilon_n \to 0$ as n increases. By using the Markov structure of the fading parallel relay network, the mutual information in (53) can be simplified as

$$
\begin{aligned}
&I\left(X, X(\mathcal{S}); Y(\mathcal{S}^c), Y|X(\mathcal{S}^c), H^N, G^N\right)\\
&\overset{(a)}{=} I\left(X; Y(\mathcal{S}^c)|X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad + I\left(X(\mathcal{S}); Y(\mathcal{S}^c)|X, X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad + I\left(X(\mathcal{S}); Y|Y(\mathcal{S}^c), X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad + I\left(X; Y|Y(\mathcal{S}^c), X^N, H^N, G^N\right)\\
&\overset{(b)}{=} I\left(X; Y(\mathcal{S}^c)|X(\mathcal{S}^c), H,^N G^N\right)\\
&\quad + I\left(X(\mathcal{S}); Y|Y(\mathcal{S}^c), X(\mathcal{S}^c), H^N, G^N\right)\\
&\overset{(c)}{\le} I\left(X; Y(\mathcal{S}^c)|H^N, G^N\right) + I(X(\mathcal{S}); Y|X(\mathcal{S}^c), H^N, G^N)
\end{aligned}
\tag{54}
$$

where (a) follows from the chain rule of mutual information, (b) follows from the Markov relations $X(\mathcal{S}) \to (X, X(\mathcal{S}^c), H^N, G^N) \to Y(\mathcal{S}^c)$ and $X \to (Y(\mathcal{S}^c), X^N, H^N, G^N) \to Y$, and (c) follows since

$$
\begin{aligned}
&I\left(X; Y(\mathcal{S}^c)|X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad = H\left(Y(\mathcal{S}^c)|X(\mathcal{S}^c), H^N, G^N\right) - H\left(Y(\mathcal{S}^c)|X, H^N, G^N\right)\\
&\quad \le I\left(X; Y(\mathcal{S}^c)|H^N, G^N\right)
\end{aligned}
\tag{55}
$$

and

$$
\begin{aligned}
&I\left(X(\mathcal{S}); Y|Y(\mathcal{S}^c), X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad = H\left(Y|Y(\mathcal{S}^c), X(\mathcal{S}^c), H^N, G^N\right) - H(Y|X^N, H^N, G^N)\\
&\quad \le I\left(X(\mathcal{S}); Y|X(\mathcal{S}^c), H^N, G^N\right).
\end{aligned}
\tag{56}
$$

The inequalities (55) and (56) are due to the Markov relations $X(\mathcal{S}^c) \to (X, H^N, G^N) \to Y(\mathcal{S}^c)$ and $X(\mathcal{S}^c) \to (X^N, H^N, G^N) \to Y$, and that conditioning reduces entropy.

Then

$$
\begin{aligned}
&I\left(X; Y(\mathcal{S}^c)|H^N, G^N\right)\\
&\quad = H\left(Y(\mathcal{S}^c)|H^N, G^N\right) - H\left(Y(\mathcal{S}^c)|X, H^N, G^N\right)\\
&\quad \le \mathsf{E}\left[\log\left(\pi e\left(1 + \sum_{k\in\mathcal{S}^c}|H_k|^2 P\right)\right)\right] - \log(\pi e)\\
&\quad = \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c}|H_k|^2 P\right)\right]
\end{aligned}
\tag{57}
$$

where the inequality follows since $X \sim \mathcal{N}_{\mathbb{C}}(0, P)$ maximizes $H(Y(\mathcal{S}^c)|H^N, G^N)$ [30]. Finally,

$$
\begin{aligned}
&I\left(X(\mathcal{S}); Y|X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad = I\left(X(\mathcal{S}); Y'|X(\mathcal{S}^c), H^N, G^N\right)\\
&\quad \le I\left(X(\mathcal{S}); Y'|H^N, G^N\right)\\
&\quad \le \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}}|G_k|^2\frac{P_r}{N}\right)\right]
\end{aligned}
\tag{58}
$$

where $Y' = \sum_{k\in\mathcal{S}} G_k X_k + Z$, and the last inequality is from the fact that a jointly Gaussian input with a diagonal covariance matrix maximizes the multiple input single output channel with per antenna power constraint [31]. Therefore, from (54) to (58), we have

$$C \le \min_{\mathcal{S}\subseteq[1:N]} \mathsf{E}\left[\mathsf{C}\left(\sum_{k\in\mathcal{S}^c}|H_k|^2 P\right) + \mathsf{C}\left(\sum_{k\in\mathcal{S}}|G_k|^2\frac{P_r}{N}\right)\right]. \tag{59}$$

REFERENCES

[1] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[2] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[3] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[5] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[6] M. H. Yassaee and M. R. Aref, "Slepian-Wolf coding over cooperative relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3462–3482, Jun. 2011.

[7] J. Hou and G. Kramer, "Short message noisy network coding with a decode—Forward option," to be published. [Online]. Available: http://arxiv.org/abs/1304.1692/

[8] X. Wu and L.-L. Xie, "On the optimal compressions in the compress-and-forward relay schemes," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, May 2013.

[9] A. Avestimehr, S. Diggvi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[10] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[11] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[12] S. Katti, I. Marić, A. Goldsmith, D. Katabi, and M. Médard, "Joint relaying and network coding in wireless networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1–5.

[13] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. ACM SIGCOMM*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[14] I. Marić, A. Goldsmith, and M. Médard, "Multihop analog network coding via amplify-and-forward: The high-SNR regime," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 793–803, Feb. 2012.

[15] U. Niesen and S. Diggavi, "The approximate capacity of the Gaussian *N*-relay diamond network," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 845–859, Feb. 2013.

[16] B. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, 2001.

[17] Y. Kochman, A. Khina, U. Erez, and R. Zamir, "Rematch-and-forward: Joint source—Channel coding for parallel relaying with spectral mismatch," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 605–622, Jan. 2014.

[18] S. S. C. Rezaei, S. O. Gharan, and A. K. Khandani, "A new achievable rate for the Gaussian parallel relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. 2009, pp. 195–198.

[19] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[20] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[21] A. S. Avestimehr and D. N. C. Tse, "Outage capacity of the fading relay channel in the low-SNR regime," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1401–1415, Apr. 2007.

[22] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[23] T. W. Ban, B. C. Jung, D. K. Sung, and W. Choi, "Performance analysis of two relay selection schemes for cooperative diversity," in *Proc. IEEE PIMRC*, Athens, Greece, Sep. 2007, pp. 1–5.

[24] M. Chen, S. Serbetli, and A. Yener, "Distributed power allocation strategies for parallel relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 552–561, Feb. 2008.

[25] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[26] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1668–1672.

[27] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[28] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[29] A. El Gamal, "On information flow in relay networks," in *Proc. IEEE Nat. Telecommun. Conf.*, Miami, FL, USA, Nov. 1981, pp. 1–5.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[31] M. Vu, "MISO capacity with per-antenna power constraint," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1268–1274, Mar. 2011.

Sang-Woon Jeon (S'07–M'11) received the B.S. and M.S. degrees in electrical engineering from Yonsei University, Seoul, Korea, in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from KAIST, Deajeon, Korea, in 2011. He is an Assistant Professor in the Department of Information and Communication Engineering at Andong National University since 2013. From 2011 to 2013, he was a Postdoctoral Associate in the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale, Lausanne, Switzerland (EPFL). His research interests include network information theory and its application to wireless communications.

Dr. Jeon won the Best Thesis Award from the EE Department, KAIST in 2012 and the ICC2015 Best Paper Award from the IEEE Communications Society in 2015.

Sung Hoon Lim (S'08–M'12) received the B.S. degree (with honors) in electrical and computer engineering from Korea University, Korea, in 2005, and the M.S. and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 2007 and 2011, respectively. From March 2012 to May 2014, he was with Samsung Electronics. He is currently a Postdoctoral Associate in the School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland. His research interests are in information theory, communication systems, and data compression.

Dr. Lim was the recipient of the Bronze Prize of the Intel Student Paper Awards in 2005 and the Gold Prize of the Samsung Humantech Paper Awards in 2011.

Bang Chul Jung (S'02–M'08–SM'14) received the B.S. degree in electronics engineering from Ajou University, Suwon, Korea, in 2002 and the M.S. and Ph.D. degrees in electrical & computer engineering from KAIST, Daejeon, Korea, in 2004 and 2008, respectively. He was a Senior Researcher/Research Professor with KAIST Institute for Information Technology Convergence, Daejeon, Korea, from January 2009 to February 2010. From March 2010 to August 2015, he was a Faculty of Gyeongsang National University. He is currently an Associate Professor of the Department of Electronics Engineering, Chungnam National University, Daejon, Korea. His research interests include 5G mobile communication systems, statistical signal processing, opportunistic communications, compressed sensing, interference management, interference alignment, random access, relaying techniques, device-to-device networks, in-network computation, and network coding.

Dr. Jung was the recipient of the Fifth IEEE Communication Society Asia-Pacific Outstanding Young Researcher Award in 2011. He was also the recipient of the Bronze Prize of Intel Student Paper Contest in 2005, the First Prize of KAIST's Invention Idea Contest in 2008, the Bronze Prize of Samsung Humantech Paper Contest in 2009, the Outstanding Research Award of Institute of Marine Industry in Gyeongsang National University in 2013, the Gaechuck Award for Excellence in Teaching of Gyeongsang National University in 2014, and the Outstanding Paper Award in Spring Conference of Korea Institute of Information and Communication Engineering in 2015.