New Upper Bounds on the Capacity of Primitive Diamond Relay Channels

Xiugang Wu
University of Delaware
Email: xwu@udel.edu

Ayfer Ozgur
Stanford University
Email: aozgur@stanford.edu

Michael Peleg and Shlomo Shamai (Shitz)
Technion-Israel Institute of Technology
Email: peleg.michael@gmail.com; sshlomo@ee.technion.ac.il

***Abstract*—Consider a primitive diamond relay channel, where a source X wants to send information to a destination with the help of two relays Y_1 and Y_2, and the two relays can communicate to the destination via error-free digital links of capacities C_1 and C_2 respectively, while Y_1 and Y_2 are conditionally independent given X. In this paper, we develop new upper bounds on the capacity of such primitive diamond relay channels that are tighter than the cut-set bound. Our results include both the Gaussian and the discrete memoryless case and build on the information inequalities recently developed in [6]–[8] that characterize the tension between information measures in a certain Markov chain.**

I. Introduction

The diamond relay channel was first introduced in [1]. It models a communication scenario where a source wants to send information to a destination with the help of two relay nodes. The channels between the source and the two relays form a broadcast channel in the first stage and the channels between the two relays and the destination form a multiple access channel in the second stage. The capacity of the diamond relay channel remains open in general. Lower bounds and upper bounds on the capacity have been developed in [1]–[3].

In this paper, we study a class of diamond relay channels where the multiple access channel between the two relays and the destination is modeled by two orthogonal links with finite capacities. This model has been of significant recent interest as it captures the (single-user) uplink communication in Cloud Radio Access Networks (C-RANs) [4]. In this paper, we will call this channel model the primitive diamond relay channel. The only known upper bound on the capacity of this channel in the current literature is the so-called cut-set bound [5]. In this paper, we develop new upper bounds on the capacity of the primitive diamond relay channel that are tighter than the cut-set bound.

Our results include both the Gaussian case and the general discrete memoryless case. In the Gaussian case, our new bound builds on the information inequality recently developed in [6]–[7].[1] This inequality characterizes the tension between information measures in a certain Markov chain, and interestingly, was proved based on using an extended isoperimetric inequality to study the geometric relations between the typical sets of the random variables involved. In the discrete memoryless case, we build on an analogous information inequality developed in [8], which was proved via using measure concentration to study the geometric relations between the typical sets. However, since the inequality in [8] only applies to a symmetric Markov chain, we will use channel simulation ideas [11]–[12] to connect the general asymmetric diamond relay channel to a symmetric channel.

The work of A. Ozgur was supported in part by NSF award CCF-1704624 and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370. The work of M. Peleg and S. Shamai was supported by the European Union's Horizon 2020 Research And Innovation Programme, grant agreement no. 694630.

[1] See also [9]–[10] for a weak version of the information inequality.

II. Channel Model

Consider a general discrete memoryless primitive diamond relay channel as depicted in Fig. 1. The source's input X is received by the two relays Y_1 and Y_2 through a broadcast channel

$$(\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$$

where $\mathcal{X}, \mathcal{Y}_1$ and $\mathcal{Y}_2$ are finite sets denoting the alphabets of the source, the relay 1 and the relay 2, respectively, and $p(y_1, y_2|x)$ is the channel transition probability; the relay 1 and the relay 2 can communicate to the destination via error-free digital links of capacities C_1 and C_2, respectively. We assume that the two relays' observations Y_1 and Y_2 are conditionally independent given X, i.e. $p(y_1, y_2|x) = p(y_1|x)p(y_2|x)$.

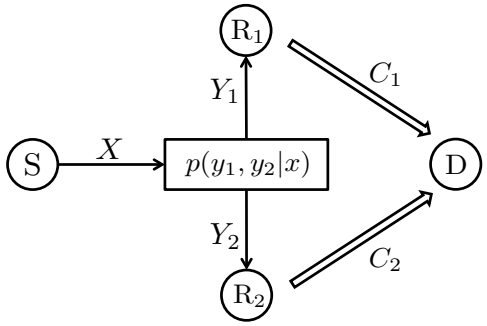

Fig. 1. Primitive diamond relay channel.

For this channel, a code of rate R for n channel uses, denoted by $(\mathcal{C}_{(n,R)}, (f_{1n}, f_{2n}), g_n)$, consists of the following:

1) A codebook at the source,

$$\mathcal{C}_{(n,R)} = \left\{x^n(m) \in \mathcal{X}^n : m \in [1 : 2^{nR}]\right\};$$

978-1-5386-6900-6/19/$31.00 ©2019 IEEE

2) Two encoding functions f_{1n} and f_{2n} at the relay 1 and relay 2 respectively,

$$f_{1n} : \mathcal{Y}_1^n \to [1 : 2^{nC_1}],$$
$$f_{2n} : \mathcal{Y}_2^n \to [1 : 2^{nC_2}];$$

3) A decoding function at the destination,

$$g_n : [1 : 2^{nC_1}] \times [1 : 2^{nC_2}] \to [1 : 2^{nR}].$$

The average probability of error of the code is defined as

$$P_e^{(n)} = \Pr(g_n(f_{1n}(Y_1^n), f_{2n}(Y_2^n)) \neq M),$$

where the message M is assumed to be uniformly drawn from the message set $[1 : 2^{nR}]$. A rate R is said to be achievable if there exists a sequence of codes $\{(\mathcal{C}_{(n,R)}, (f_{1n}, f_{2n}), g_n)\}_{n=1}^{\infty}$ such that the average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. The capacity of the primitive diamond relay channel is the supremum of all achievable rates, denoted by $C(C_1, C_2)$.

A. Gaussian Primitive Diamond Relay Channel

In this paper, we will also be interested in the Gaussian version of this channel as depicted in Fig. 2. Here, we have

$$\begin{cases} Y_1 = X + W_1 \\ Y_2 = X + W_2 \end{cases}$$

where $X \in \mathbb{R}$ is the source signal constrained to average power P, and $W_1 \sim \mathcal{N}(0, N_1), W_2 \sim \mathcal{N}(0, N_2)$ are Gaussian noises that are independent of each other and X.

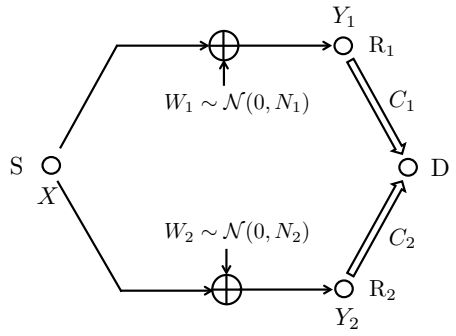

Fig. 2. Gaussian primitive diamond relay channel.

III. Main Results

The only known upper bound on the capacity of the primitive diamond relay channel is given by the cut-set bound.

Proposition 3.1 (Cut-set Bound): For a primitive diamond relay channel depicted in Fig. 1, if a rate R is achievable, then there exists some $p(x)$ such that

$$\begin{cases} R \leq I(X; Y_1, Y_2) & (1) \\ R \leq I(X; Y_1) + C_2 & (2) \\ R \leq I(X; Y_2) + C_1 & (3) \\ R \leq C_1 + C_2. & (4) \end{cases}$$

The main results of this paper are to develop new upper bounds on the capacity of the primitive diamond relay channel that are tighter than the cut-set bound. In the following, we present our new upper bounds for the Gaussian case and the discrete memoryless case, respectively. The proofs of these bounds are provided in Sections IV and V.

A. Gaussian Case

Our new upper bound in the Gaussian case is given by the following theorem.

Theorem 3.1: For a Gaussian primitive diamond relay channel depicted in Fig. 2, if a rate R is achievable, then there exists $\theta_i \in [\arcsin(2^{-C_i}), \pi/2], i \in \{1, 2\}$ such that

$$\begin{cases} R \leq 1/2 \log(1 + \mathbf{snr}_1) + C_2 + \log \sin \theta_2 & (5) \\ R \leq 1/2 \log(1 + \mathbf{snr}_1) + \min\limits_{\omega_2 \in \left(\frac{\pi}{2} - \theta_2, \frac{\pi}{2}\right]} h_2(\omega_2; \theta_2) & (6) \\ R \leq 1/2 \log(1 + \mathbf{snr}_2) + C_1 + \log \sin \theta_1 & (7) \\ R \leq 1/2 \log(1 + \mathbf{snr}_2) + \min\limits_{\omega_1 \in \left(\frac{\pi}{2} - \theta_1, \frac{\pi}{2}\right]} h_1(\omega_1; \theta_1) & (8) \\ R \leq C_1 + C_2 + \log \sin \theta_1 + \log \sin \theta_2 & (9) \end{cases}$$

where $\mathbf{snr}_i = P/N_i$ for $i \in \{1, 2\}$, and $h_i(\omega; \theta)$ is defined as

$$\frac{1}{2} \log \left(\frac{\left[P(N_1 + N_2) + N_1 N_2 \sin^2 \omega - 2P\sqrt{N_1 N_2} \cos \omega\right] \sin^2 \theta}{N_i (P + N_{\bar{i}})(\sin^2 \theta - \cos^2 \omega)} \right)$$

for $i, \bar{i} \in \{1, 2\}$, $i \neq \bar{i}$.

One can show that the above bound is tighter than the cut-set bound. In particular, when specialized to the Gaussian case, the cut-set bound in Proposition 3.1 says that any achievable rate R satisfies

$$\begin{cases} R \leq 1/2 \log(1 + \mathbf{snr}_1 + \mathbf{snr}_2) & (10) \\ R \leq 1/2 \log(1 + \mathbf{snr}_1) + C_2 & (11) \\ R \leq 1/2 \log(1 + \mathbf{snr}_2) + C_1 & (12) \\ R \leq C_1 + C_2. & (13) \end{cases}$$

Our bound (5)–(9) is in general tighter than the cut-set bound (10)–(13) because of the following:

1) The constraints (6) and (8) in our bound imply the broadcast bound (10) in the cut-set bound. Indeed, in constraint (6), for any $\theta_2 \in [\arcsin(2^{-C_2}), \pi/2]$,

$$\min_{\omega_2 \in \left(\frac{\pi}{2} - \theta_2, \frac{\pi}{2}\right]} h_2(\omega_2; \theta_2) \leq h_2(\pi/2; \theta_2)$$
$$= 1/2 \log(1 + \mathbf{snr}_1 + \mathbf{snr}_2) - 1/2 \log(1 + \mathbf{snr}_1)$$

and therefore constraint (6) implies (10); similarly, constraint (8) also implies (10).

2) Since $\log \sin \theta_i \leq 0$ for any $\theta_i \in [\arcsin(2^{-C_i}), \pi/2]$, $i \in \{1, 2\}$, the constraints (5), (7) and (9) in our bound are in general stricter than (11), (12) and (13) in the cut-set bound.

Now, to see that our bound can be indeed *strictly* tighter than the cut-set bound, consider the special symmetric case when $\mathbf{snr}_1 = \mathbf{snr}_2$. In this case, the optimal θ_1 and θ_2 in our bound have to be strictly less than $\pi/2$ because otherwise the second term in both (6) and (8) can be made arbitrarily small by taking ω_2 and ω_1 to be arbitrarily close to zero, thereby constraining the achievable rate R by $\frac{1}{2} \log(1 + \mathbf{snr}_1)$ and $\frac{1}{2} \log(1 + \mathbf{snr}_2)$, respectively. This implies that when $\mathbf{snr}_1 = \mathbf{snr}_2$, the constraints (5), (7) and (9) are strictly stricter than (11), (12) and (13) respectively, and therefore our bound is strictly tighter than the cut-set bound.

In Fig. 3 we plot the upper bound in Theorem 3.1 together with the cut-set bound for the symmetric case when $\mathbf{snr}_1 = \mathbf{snr}_2 =: \mathbf{snr}_0$ and $C_1 = C_2 =: C_0$, under two different values of $\mathbf{snr}_0$. For reference, we also plot the rate achieved by time sharing between decode-and-forward and a compress-and-forward with Gaussian input distribution and Gaussian quantization [13]. Note that all the three bounds coincide in the regime when C_0 is relatively small. Indeed, when $1/2\log(1+\mathbf{snr}_0) \geq 2C_0$, both relays can decode the transmitted message, and the cut-set bound, which is equal to the multiple-access bound $2C_0$ in this regime, can be trivially achieved. Known achievable schemes fail to achieve the multiple-access bound when $1/2\log(1+\mathbf{snr}_0) < 2C_0$. As argued above, our new bound establishes that the capacity itself deviates from the multiple-access bound when $1/2\log(1+\mathbf{snr}_0) < 2C_0$ and therefore the multiple-access bound is achievable if and only if $1/2\log(1+\mathbf{snr}_0) \geq 2C_0$. Also note that from these figures one can visually observe that the new upper bound reaches the value $C(\infty,\infty) = 1/2\log(1+\mathbf{snr}_1+\mathbf{snr}_2)$ only as $C_0 \to \infty$. Indeed, one can formalize this and show the following corollary under the more general asymmetric setup.

Corollary 3.1: For a Gaussian primitive diamond relay channel depicted in Fig. 2, its capacity $C(C_1, C_2)$ is bounded away from $C(\infty,\infty)$ if C_1 or C_2 is finite.

Note that an immediate implication of the above corollary is to say that in order to achieve the broadcast bound $C(\infty,\infty)$, both C_1 and C_2 have to be infinity. Proof of this corollary follows along similar lines as the proof of Theorem 1 in [7] and is omitted in this paper.

B. Discrete Memoryless Case

In the discrete memoryless case, our new bound is given by the following theorem.

Theorem 3.2: For a discrete memoryless primitive diamond relay channel depicted in Fig. 1, if a rate R is achievable, then there exists some $p(x)$ and $a_1, a_2 \geq 0$ such that

$$\begin{cases} R \leq I(X;Y_1,Y_2) & (14)\\ R \leq I(X;Y_1) + C_2 - a_2 & (15)\\ R \leq I(X;Y_1,\tilde{Y}_2) + H_2\left(\sqrt{a_2 \ln 2/2}\right) \\ \qquad + \sqrt{a_2 \ln 2/2}\log(|\mathcal{Y}_2|-1) - a_2 & (16)\\ R \leq I(X;Y_2) + C_1 - a_1 & (17)\\ R \leq I(X;\tilde{Y}_1,Y_2) + H_2\left(\sqrt{a_1 \ln 2/2}\right) \\ \qquad + \sqrt{a_1 \ln 2/2}\log(|\mathcal{Y}_1|-1) - a_1 & (18)\\ R \leq C_1 + C_2 - a_1 - a_2 & (19) \end{cases}$$

for any random variables $\tilde{Y}_1$ and $\tilde{Y}_2$ with the same conditional distributions as Y_1 and Y_2 given X, i.e., $p_{\tilde{Y}_1|X}(y_1|x) = p_{Y_1|X}(y_1|x)$ and $p_{\tilde{Y}_2|X}(y_2|x) = p_{Y_2|X}(y_2|x)$, where $H_2(\cdot)$ is defined as $H_2(r) = -r\log r - (1-r)\log(1-r)$ for $r \in [0,1]$ and $H_2(r) = 0$ for $r \notin [0,1]$.

The evaluation of the bound in Theorem 3.2 involves optimizing over all the $\tilde{Y}_1$ and $\tilde{Y}_2$ random variables that have the same conditional distributions as Y_1 and Y_2 given X. While this optimization may not be trivial in general, note that any valid choice of $\tilde{Y}_1$ and $\tilde{Y}_2$ would provide an upper bound on the capacity. Moreover, in the case when the broadcast channel is symmetric or stochastically degraded, it is straightforward to determine the optimal $\tilde{Y}_1$ and $\tilde{Y}_2$:

1) When the broadcast channel is symmetric, i.e. $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}$ and $p_{Y_1|X}(y|x) = p_{Y_2|X}(y|x)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, choosing $\tilde{Y}_1 = Y_2$ and $\tilde{Y}_2 = Y_1$ will minimize the mutual information term in constraints (18) and (16) respectively and this gives the tightest bound.
2) Now consider the case when the broadcast channel is stochastically degraded, i.e. there exists some channel transition probability $q(y_2|y_1)$ such that $p(y_2|x) = \sum_{y_1} p(y_1|x)q(y_2|y_1)$. In this case, the optimal $\tilde{Y}_1$ is such that $p_{\tilde{Y}_1,Y_2|X}(y_1,y_2|x) = p_{Y_1|X}(y_1|x)q(y_2|y_1)$, i.e. $X - \tilde{Y}_1 - Y_2$ form a Markov chain; this is because in constraint (18), $I(X;\tilde{Y}_1,Y_2) \geq I(X;\tilde{Y}_1) = I(X;Y_1)$, where the inequality holds with equality if and only if $X - \tilde{Y}_1 - Y_2$ form a Markov chain. Similarly, the optimal $\tilde{Y}_2$ is such that $p_{Y_1,\tilde{Y}_2|X}(y_1,y_2|x) = p_{Y_1|X}(y_1|x)q(y_2|y_1)$, i.e. $X - Y_1 - \tilde{Y}_2$ form a Markov chain; with such $\tilde{Y}_2$, $I(X;Y_1,\tilde{Y}_2)$ in constraint (16) achieves the minimal value $I(X;Y_1)$.

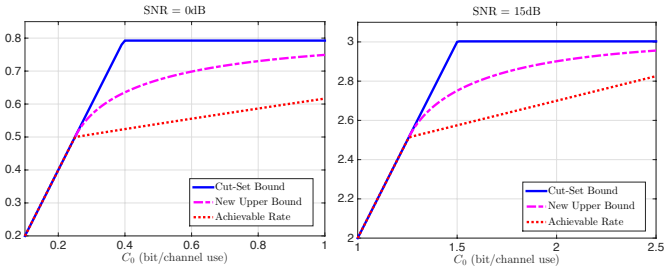

Fig. 3. Capacity bounds for the Gaussian primitive diamond relay channel.

It is easy to observe that the above bound is tighter than the cut-set bound (1)–(4). In particular, since $a_1, a_2 \geq 0$ in Theorem 3.2, constraints (15), (17) and (19) are in general tighter than (2), (3) and (4) respectively, and therefore our bound in Theorem 3.2 is in general tighter than the cut-set bound. To see that our bound can be indeed *strictly* tighter, consider the above mentioned case when the broadcast channel is symmetric or stochastically degraded. In this case, the optimal a_1 and a_2 in our bound have to be strictly positive because otherwise bounds (16) and (18) would constrain the rate R by $I(X;Y_1)$. This implies that when the broadcast channel is symmetric or stochastically degraded, constraints (15), (17) and (19) are strictly tighter than (2), (3) and (4) respectively, and therefore our bound is strictly tighter than the cut-set bound.

IV. Gaussian Primitive Diamond Relay Channel

Our new upper bound for the Gaussian primitive diamond relay channel, i.e. Theorem 3.1, builds on the following lemma proved in [6]–[7].

Lemma 4.1: Consider a Markov chain $J_n - U^n - X^n - V^n$, where X^n, U^n and V^n are n-length random vectors and $J_n = f_n(U^n)$ is a deterministic mapping of U^n to a set of integers. Assume moreover that U^n and V^n are independent white Gaussian vectors given X^n such that $U^n \sim \mathcal{N}(X^n, N_u\, I_{n\times n})$ and $V^n \sim \mathcal{N}(X^n, N_v\, I_{n\times n})$ where $I_{n\times n}$ denotes the identity matrix, and $\mathbb{E}[\|X^n\|^2] = nP$, and $H(J_n|X^n) = -n \log \sin\theta_n$ for some $\theta_n \in [0, \pi/2]$. Then the inequality (20) at the top of the next page holds for any n.

Note that the above lemma provides an upper bound on $H(J_n|V^n)$ in terms of $H(J_n|X^n)$ for a Markov chain that satisfies the conditions of the lemma. The proof of the lemma can be found in [6]–[7] and is omitted here. Interestingly, such a proof is based on considering the i.i.d. extensions of the random variables (X^n, U^n, V^n, J_n) and studying the geometric relations between their typical sets using an extended isoperimetric inequality. We now use this lemma to prove Theorem 3.1.

A. Proof of Theorem 3.1

Suppose a rate R is achievable. Then there exists a sequence of $(\mathcal{C}_{(n,R)}, (f_{1n}, f_{2n}), g_n)$ codes such that the average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. Let the transmissions by the relay 1 and relay 2 be denoted by $J_{1n} = f_{1n}(Y_1^n)$ and $J_{2n} = f_{2n}(Y_2^n)$ respectively. By standard information theoretic arguments, for this sequence of codes we have

$$\begin{aligned}
nR &= H(M) \\
&= I(M; J_{1n}, J_{2n}) + H(M|J_{1n}, J_{2n}) \\
&\le I(M; J_{1n}, J_{2n}) + n\mu \qquad (21)\\
&\le I(X^n; Y_1^n, J_{2n}) + n\mu \\
&= I(X^n; Y_1^n) + I(X^n; J_{2n}|Y_1^n) + n\mu \qquad (22)\\
&\le n/2 \log(1 + \mathbf{snr}_1) + H(J_{2n}|Y_1^n) - H(J_{2n}|X^n) + n\mu, \qquad (23)
\end{aligned}$$

for any $\mu > 0$ and n sufficiently large.

Given (23), the standard way to proceed would be to upper bound the first entropy term by $H(J_{2n}|Y_1^n) \le H(J_{2n}) \le nC_2$ and lower bound the second entropy term $H(J_{2n}|X^n)$ simply by 0. This would lead to the cross-cut bound (11) in the cut-set bound. However, as pointed out in [6]–[7], this bounding procedure will result in a loose bound since it does not capture the inherent tension between how large the first entropy term can be and how small the second one can be. Instead, we can use Lemma 4.1 to more tightly upper bound the difference $H(J_{2n}|Y_1^n) - H(J_{2n}|X^n)$ in (23). In particular, we assume that $H(J_{2n}|X^n) = -n \log \sin\theta_{2n}$ for some $\theta_{2n} \in [\arcsin(2^{-C_2}), \pi/2]$, and leave it as it is in (23), yielding

$$R \le 1/2 \log(1 + \mathbf{snr}_1) + C_2 + \log \sin\theta_{2n} + \mu,$$

which is bound (5). Then applying Lemma 4.1 to the Markov chain $J_{2n} - Y_2^n - X^n - Y_1^n$, we have

$$H(J_{2n}|Y_1^n) - H(J_{2n}|X^n) \le n \cdot \min_{\omega_2 \in \left(\frac{\pi}{2} - \theta_{2n}, \frac{\pi}{2}\right]} h_2(\omega_2; \theta_{2n}).$$

Plugging this into (23), we conclude that

$$R \le 1/2 \log(1 + \mathbf{snr}_1) + \min_{\omega_2 \in \left(\frac{\pi}{2} - \theta_{2n}, \frac{\pi}{2}\right]} h_2(\omega_2; \theta_{2n}) + \mu,$$

which proves bound (6). By symmetry, one can prove bounds (7) and (8) similarly.

Finally, continuing with (21) we have

$$\begin{aligned}
nR &\le I(M; J_{1n}, J_{2n}) + n\mu \\
&\le I(X^n; J_{1n}, J_{2n}) + n\mu \\
&= I(X^n; J_{1n}) + I(X^n; J_{2n}|J_{1n}) + n\mu \\
&= H(J_{1n}) - H(J_{1n}|X^n) \\
&\quad + H(J_{2n}|J_{1n}) - H(J_{2n}|J_{1n}, X^n) + n\mu \\
&\le H(J_{1n}) - H(J_{1n}|X^n) + H(J_{2n}) - H(J_{2n}|X^n) + n\mu \\
&\le nC_1 + nC_2 + n \log \sin\theta_{1n} + n \log \sin\theta_{2n} + n\mu.
\end{aligned}$$

This proves bound (9) and concludes the proof of the theorem.

V. Discrete Memoryless Primitive Diamond Relay Channel

In the discrete memoryless case, we build on the following lemma that was proved in [8] via measure concentration.

Lemma 5.1: Consider a Markov chain $J_n - U^n - X^n - V^n$, where X^n, U^n and V^n are n-length discrete random vectors and $J_n = f_n(U^n)$ is a deterministic mapping of U^n to a set of integers. Assume moreover that U^n and V^n are memoryless and conditionally i.i.d. given X^n, i.e.,

1) $p(u^n, v^n|x^n) = \prod_{i=1}^n p(u_i|x_i)p(v_i|x_i)$;
2) $p_{U|X}(t|x) = p_{V|X}(t|x)$, for any $x \in \mathcal{X}$ and $t \in \mathcal{T}$, where $\mathcal{X}$ and $\mathcal{T}$ denote the alphabet of X and the common alphabet of U and V respectively.

Suppose $H(J_n|X^n) = na_n$ for some $a_n \ge 0$. Then

$$H(J_n|V^n) \le n\left[H_2\left(\sqrt{\frac{a_n \ln 2}{2}}\right) + \sqrt{\frac{a_n \ln 2}{2}} \log(|\mathcal{T}| - 1)\right] \qquad (24)$$

Analogous to Lemma 4.1 in the Gaussian case, Lemma 5.1 provides an upper bound on $H(J_n|V^n)$ in terms of $H(J_n|X^n)$ in the discrete case. However, note that applying this lemma requires a symmetric Markov chain structure, which may not be always satisfied under a general discrete memoryless diamond relay setup. For this, in the following we will combine channel simulation ideas with the above lemma to prove Theorem 3.2.

A. Proof of Theorem 3.2

Suppose a rate R is achievable. Then along the same lines to reach (22), we have

$$nR \le I(X^n; Y_1^n) + H(J_{2n}|Y_1^n) - H(J_{2n}|X^n) + n\mu, \quad (25)$$

for any $\mu > 0$ and n sufficiently large. We now fix $H(J_{2n}|X^n) = na_n$ and leave it as it is in the above inequality, yielding

$$nR \le I(X^n; Y_1^n) + nC_2 - na_n + n\mu, \qquad (26)$$

$$H(J_n|V^n) \le n \cdot \min_{\omega \in \left(\frac{\pi}{2}-\theta_n, \frac{\pi}{2}\right]} \frac{1}{2}\log\left(\frac{P(N_u+N_v)+N_uN_v\sin^2\omega - 2P\sqrt{N_uN_v}\cos\omega}{N_u(P+N_v)(\sin^2\theta_n - \cos^2\omega)}\right) \tag{20}$$

which proves bound (15).

To prove bound (16), we will upper bound $H(J_{2n}|Y_1^n)$ in terms of a_n using Lemma 5.1. For this, we will first resort to channel simulation theory [11]–[12] to construct an auxiliary random variable $\tilde{Y}_2^n$ so that $\tilde{Y}_2^n$ and Y_2^n are conditionally i.i.d. given X^n. Specifically, consider the channel simulation setup as depicted in Fig. 4, where we want to simulate some channel $p_{\tilde{Y}_2|X}(y_2|x)$ such that $p_{\tilde{Y}_2|X}(y_2|x) = p_{Y_2|X}(y_2|x)$, i.e., $\tilde{Y}_2$ has the same conditional distribution as Y_2. The simulation encoder sees the source X^n, side information Y_1^n, and a common random variable K_n which is uniformly distributed on $[1 : 2^{nR_2}]$ and independent of those random variables $(X^n, Y_1^n, Y_2^n, J_{2n})$ associated with the original channel, and it generates a simulation codeword $S_n \in [1 : 2^{nR_1}]$ based on a randomized encoding function $E_n(X^n, Y_1^n, K_n)$. The simulation decoder also observes Y_1^n and K_n, and upon receiving S_n it outputs a random variable $\tilde{Y}_2^n$ based on a randomized decoding function $D_n(S_n, Y_1^n, K_n)$.

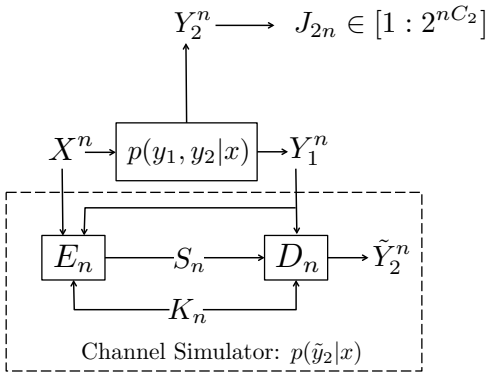

Fig. 4. Channel simulation.

Following the similar lines as in [11]–[12] and [14], it can be shown that the channel $p_{\tilde{Y}_2|X}(\tilde{y}_2|x)$ can be simulated in the above setup, i.e., $\tilde{Y}_2^n$ is (essentially) the same as if it is generated by passing X^n through the channel $p_{\tilde{Y}_2|X}(\tilde{y}_2|x)$, if $R_1 = I(X;\tilde{Y}_2|Y_1)+\epsilon$ and R_2 is sufficiently large. In this case, $\tilde{Y}_2^n$ and Y_2^n are conditionally identically distributed given X^n, and due to the Markov chain $Y_2^n - X^n - (Y_1^n, S_n, K_n) - \tilde{Y}_2^n$ they are also conditionally independent given X^n. Therefore one can apply Lemma 5.1 to the Markov chain $J_{2n} - Y_2^n - X^n - \tilde{Y}_2^n$ and obtain that $H(J_{2n}|\tilde{Y}_2^n) \le$ R.H.S. of (24) with $\mathcal{T}$ replaced by $\mathcal{Y}_2$.

Now consider expanding $H(J_{2n}, S_n, K_n|Y_1^n)$ in two different ways as follows:

$$\begin{aligned}
&H(J_{2n}, S_n, K_n|Y_1^n)\\
&= H(J_{2n}|Y_1^n) + H(K_n|Y_1^n, J_{2n}) + H(S_n|Y_1^n, J_{2n}, K_n)\\
&= H(K_n|Y_1^n) + H(S_n|Y_1^n, K_n) + H(J_{2n}|Y_1^n, K_n, S_n).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&H(J_{2n}|Y_1^n)\\
&= H(K_n|Y_1^n) + H(S_n|Y_1^n, K_n) + H(J_{2n}|Y_1^n, K_n, S_n)\\
&\quad - H(K_n|Y_1^n, J_{2n}) - H(S_n|Y_1^n, J_{2n}, K_n)\\
&\le H(K_n|Y_1^n) + H(S_n|Y_1^n, K_n) + H(J_{2n}|Y_1^n, K_n, S_n)\\
&\quad - H(K_n|Y_1^n, J_{2n})\\
&= H(S_n|Y_1^n, K_n) + H(J_{2n}|Y_1^n, K_n, S_n) &(27)\\
&= H(S_n|Y_1^n, K_n) + H(J_{2n}|Y_1^n, K_n, S_n, \tilde{Y}_2^n) &(28)\\
&\le n(I(X;\tilde{Y}_2|Y_1)+\epsilon) + H(J_{2n}|\tilde{Y}_2^n) &(29)\\
&\le n(I(X;\tilde{Y}_2|Y_1)+\epsilon) + \text{R.H.S. of (24)} &(30)
\end{aligned}$$

where (27) follows because K_n is independent of (Y_1^n, J_{2n}), (28) follows from the Markov chain $J_{2n} - (Y_1^n, K_n, S_n) - \tilde{Y}_2^n$, and (29) follows because $H(S_n|Y_1^n, K_n) \le nR_1 = n(I(X;\tilde{Y}_2|Y_1)+\epsilon)$ and removing condition does not reduce entropy. Plugging (30) into (25) yields bound (16). Bounds (17) and (18) can be proved similarly. Bounds (14) and (19) are straightforward. This completes the proof of Theorem 3.2.

REFERENCES

[1] B. E. Schein. Distributed Coordination in Network Information Theory. PhD thesis, Massachusetts Institute of Technology, 2001.

[2] W. Kang, N. Liu, and W. Chong, "The Gaussian multiple access diamond channel," *IEEE Trans. Inf. Theory*, vol. 61, pp. 6049–6059, Nov. 2015.

[3] S. S. Bidokhti and G. Kramer, "Capacity bounds for diamond networks with an orthogonal broadcast channel," *IEEE Trans. Inf. Theory*, vol. 62, pp. 7103–7122, 2016.

[4] S.-H. Park, O. Simeone, O. Sahin and S. Shamai, "Fronthaul compression for cloud radio access networks: Signal processing advances inspired by network information theory," IEEE Signal Process. Mag., vol. 31, no. 6, pp. 69–79, Oct. 2014.

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

[6] X. Wu, L. Barnes and A. Ozgur, "The geometry of the relay channel," in *Proc. of IEEE Int. Symposium on Information Theory*, Aachen, Germany, June 2017.

[7] X. Wu, L. Barnes and A. Ozgur, ""The Capacity of the Relay Channel": Solution to Cover's Problem in the Gaussian Case," *IEEE Trans. Inf. Theory*, vol. 65, pp. 255–275, January 2019.

[8] X. Wu, A. Ozgur, and L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," *IEEE Trans. Inf. Theory*, vol. 63, pp. 2254–2277, April 2017.

[9] X. Wu and A. Ozgur, "Cut-set bound is loose for Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 64, pp. 1023–1037, 2018.

[10] J. Liu and A. Ozgur, "Capacity Upper Bounds for the Relay Channel via Reverse Hypercontractivity," available online at http://arxiv.org/abs/1811.11303.

[11] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[12] P. Cuff, H. Permuter and T. M. Cover, "Coordination Capacity," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4181–4206, September 2010.

[13] I. E. Aguerri, A. Zaidi, G. Caire, and S. Shamai, "On the capacity of cloud radio access networks with oblivious relaying," available online at http://arxiv.org/abs/1701.07237.

[14] F. Xue, "A new upper bound on the capacity of a primitive relay channel based on channel simulation," *IEEE Trans. Inform. Theory*, vol. 60, pp. 4786–4798, Aug. 2014.