Achievable error exponents for the two-way parallel DMC

Kenneth Palacio-Baus*†, Natasha Devroye*
*University of Illinois at Chicago, {kpalac2, devroye}@uic.edu
†University of Cuenca, Ecuador

***Abstract*—We investigate error exponent regions for the parallel two-way DMC in which each terminal sends its own message and provides feedback to the other terminal. Various error exponents are presented in different rate-region regimes based on the relative rates and zero-error capacities of both directions. The schemes employed are extensions of error exponents for one-way DMCs with noiseless, rate-limited and noisy feedback[1].**

I. Introduction

Shannon [1] introduced the two-way discrete memoryless channel (DMC), and derived inner and outer bounds to its capacity region. We focus on the two-way *parallel* DMC, whose capacity region is a rectangle determined by the one-way capacity of each link. Adaptation / interaction, or using the feedback present in two-way channels, cannot increase this capacity region, but may be exploited to attain larger error exponents. We consider that either one or both directions have a positive zero-error capacity C_0, and improve the reliability at rate-pairs in the small error regime. $C_0 > 0$ also alleviates synchronization issues in variable length coding (VLC), since the beginning / end of a message is signaled without error[2].

In the two-way setting, a terminal may transmit messages with small error at all rates below the one-way capacity, with zero-error at all rates below C_0, provide noiseless feedback for the other terminal limited to a certain rate below C_0, provide noisy feedback, or any combination of the above. We present achievable schemes and error exponents for the two-way parallel DMC based on coding schemes for the one-way DMC with feedback using VLC. The one-way reliability function is defined for VLC as:

$$E(\bar{R}) = \lim_{\mathsf{E}[\Delta]\to\infty} \frac{-1}{\mathsf{E}[\Delta]} \log \mathsf{P}_e(\bar{R}, \Delta),$$

for $0 \leq \bar{R} \leq C$ (for C the small error capacity), transmission time Δ, and probability of error $\mathsf{P}_e(\bar{R}, \Delta)$. Next we present some fundamental results:

Burnashev's reliability: for any DMC of capacity C, zero-error capacity of zero, and *noiseless output feedback* using VLC, Burnashev [4] demonstrated that $E(R) \leq E_{\text{Burn}}(\bar{R})$:

$$E_{\text{Burn}}(\bar{R}) := C_1 \left(1 - \frac{\bar{R}}{C}\right), \quad \text{for } 0 \leq \bar{R} \leq C, \qquad (1)$$

where $C_1 = \max_{x_1,x_2} D\left(p(y|x_1)||p(y|x_2)\right)$ is the Kullback-Leibler divergence of the distributions induced by the two most distinguishable symbols of the forward direction alphabet.

Yamamoto-Itoh's scheme [5]: utilizes noiseless feedback in a two-stage VLC scheme to achieve Burnashev's upper bound. In the message stage, a capacity achieving code is used to send a message whose preliminary estimate is fed back without error. In the control stage, the encoder indicates whether the receiver should accept the decision (ACK), or await a retransmission (NACK). The control message estimate is also fed back to keep synchronization. Errors result from a wrong preliminary decision and a missed NACK. Retransmissions occur if the decoder declares a NACK, which happens with an exponentially small probability and the expected number of transmissions for a message tends to one.

Forney's error exponent [6]: is attained by using a *single bit of noiseless feedback* to request a retransmission when decoding leads to an erasure. Then, $E(\bar{R}) \geq E_{\text{Forn}}\left(\bar{R}\right)$, where:

$$E_{\text{Forn}}\left(\bar{R}\right) := E_{\text{sp}}(\bar{R}) + C - \bar{R}, \text{ for } \bar{R}_\infty \leq \bar{R} \leq C. \qquad (2)$$

Above, $E_{\text{sp}}(\bar{R})$ corresponds to the sphere packing bound for a DMC without feedback, and $\bar{R}_\infty$ to the smallest rate for which the sphere packing upper bound tends to infinity [7, Sec. 5.8].

Rate-limited noiseless feedback: this interesting regime has seen limited work – [8] characterized the noiseless feedback rate needed to attain Burnashev's bound. In Section III, we extend these results to obtain achievable error exponents with noiseless rate-limited feedback.

Noisy feedback: this more complicated case, due to synchronization issues, was studied for one-way DMCs using VLC in [2], [9], and for the two-way parallel DMC in [3].

Contributions: We present achievable error exponents of the one-way DMC with limited-rate noiseless feedback in Section III. We use these results in Section IV for two-way systems, where either direction may have $C_0 > 0$. Depending on the availability and amount of C_0, a terminal may provide rate-limited noiseless feedback to the other direction in addition to the transmission of its own messages (either with zero or small error). The operating rate-pair determines if this noiseless feedback can be exploited to either exceed Forney's reliability, achieve Burnashev's bound, or attain infinite reliability. Due to space constraints, all proofs are relegated to the Appendix of the extended version of this manuscript, available at [10].

[1]The work of N. Devroye and K. Palacio-Baus was partially supported by NSF under award 1815428. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

[2]Synchronization over channels with noisy feedback has been addressed in [2], and extended to two-way channels in [3].

II. Problem statement

Let $(\mathcal{X}, W, \mathcal{Y})$ denote a DMC characterized by finite input and output alphabets $\mathcal{X}, \mathcal{Y}$ and transition probability $W(y|x)$, for the transmission of equally likely messages from set $\mathcal{M}$. Let W^n denote n uses of the channel, then $W^n(y^n|x^n) = \prod_{k=1}^n W(y_k|x_k)$ for $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$. For systems with active noiseless feedback, let $R_{\text{FB}} \in \mathbb{R}^+$ be the available rate of the feedback channel. For two-way channels, each terminal is denoted by T_i for $i = 1, 2$. Let a two-way $(\mathcal{X}_1, \mathcal{Y}_1, W(y_1 y_2|x_1 x_2), \mathcal{Y}_2, \mathcal{X}_2)$ DMC be characterized by a set of transition probability mass functions $W(y_1 y_2|x_1 x_2)$, finite input and output alphabets $\mathcal{X}_i$, $\mathcal{Y}_i$, and message sets $\mathcal{M}_1$ and $\mathcal{M}_2$. In the two-way parallel DMC, $W(y_1 y_2|x_1 x_2) = W_{12}(y_2|x_1) \cdot W_{21}(y_1|x_2)$, where subscripts denote the communication direction from T_i to T_{3-i}. This is equivalent to two independent links operating in parallel and opposite directions.

A. The one-way DMC

The concepts of small-error capacity C, and zero-error capacity C_0, for a one-way $(\mathcal{X}, W, \mathcal{Y})$ DMC were introduced by Shannon in [11] and [12] respectively.

Definition 1: A variable length block code $\mathcal{C}(\mathcal{M}, R_{\text{FB}}, N)$ for a one-way $(\mathcal{X}, W, \mathcal{Y})$ DMC with noiseless rate-limited R_{FB} active feedback and block length N, comprises:

- A set of equally likely messages $\mathcal{M}$.
- A set of forward channel encoding functions: $x_n : \mathcal{M} \times \mathcal{Z}^{n-1} \to \mathcal{X}_n$, where Z^n is the sequence received through the rate-limited noiseless feedback link to produce channel inputs $X_n = x_n(M, Z^{n-1})$.
- A set of feedback channel encoding functions: $z_n : \mathcal{Y}^{n-1} \to \mathcal{Z}$, which produce feedback inputs $Z_n = z_n(Y^{n-1})$, with $|Z^N| \leq 2^{NR_{\text{FB}}}$ per block of length N.
- A set of forward decoding functions: $\phi_n : \mathcal{Y}^n \to \mathcal{M}$,

for $n = 1, 2, ..., \Delta$, where Δ corresponds to the transmission time (a random variable), and is a stopping time for which $E[\Delta] \leq N$. Let $\bar{R} = \frac{\log|\mathcal{M}|}{E[\Delta]}$ define the average transmission rate, and let $\mathsf{P}_e(\bar{R}, \Delta, R_{\text{FB}})$ (argument R_{FB} is present according to the availability of feedback) be the maximum error probability attained among all messages at rate $\bar{R}$ and decoding (not erasure) occurring at time Δ, with a noiseless feedback of rate R_{FB}. Then, $\mathsf{P}_e(\bar{R}, \Delta, R_{\text{FB}}) = \max_{M \in \mathcal{M}} P[\phi_\Delta(Y^\Delta) \neq m | M = m \text{ sent}]$.

Definition 2: An error exponent is achievable at an expected rate $\bar{R}$ over a one-way DMC with rate-limited feedback if there exists a sequence of VLC codes such that:

$$E(\bar{R}, R_{\text{FB}}) \geq \lim_{\mathsf{E}[\Delta] \leq N, N \to \infty} \frac{-1}{\mathsf{E}[\Delta]} \log \mathsf{P}_e(\bar{R}, \Delta, R_{\text{FB}}),$$

for $C_0 < \bar{R} \leq C$. $E(\bar{R}, R_{\text{FB}}) = \infty$ for $0 \leq \bar{R} \leq C_0, \forall R_{\text{FB}}$.

B. The two-way parallel DMC

A parallel two-way DMC is formed by terminals T_i for $i = 1, 2$, and channels $\left(\mathcal{X}_i, W_{i,(3-i)}(y_{3-i}|x_i), \mathcal{Y}_{3-i}\right)$. Let $\bar{R}_{i,(3-i)}$ be the expected rate in the $i \to (3-i)$ direction, and $\Delta_{i,(3-i)}$ the transmission time[3] at which decoding decision about message M_i is made at T_{3-i}.

[3]Each direction has its own transmission time.

Definition 3: Terminals T_1 and T_2 *interact* if their corresponding channel inputs at time n adapt to past outputs as $X_{i,n} = x_{i,n}\left(M_i, y_i^{n-1}\right)$.

Definition 4: A two-way variable length code $\mathcal{C}(\mathcal{M}_1, \mathcal{M}_2, N)$ for a two-way parallel $(\mathcal{X}_1, \mathcal{Y}_1, W(y_1 y_2|x_1 x_2), \mathcal{Y}_2, \mathcal{X}_2)$ DMC comprises:

- Two sets of equally likely messages $\mathcal{M}_i$.
- Two sets of encoding functions $x_{i,n} : \mathcal{M}_i \times \mathcal{Y}_i^{n-1} \to \mathcal{X}_{i,n}$, producing channel inputs $X_{i,n} = x_{i,n}\left(M_i, Y_i^{n-1}\right)$,
- Two sets of decoding functions $\phi_{i,n} : \mathcal{Y}_i^n \to \mathcal{M}_i$,

for $n = 1, 2, ..., \Delta_{i,(3-i)}$, where $\Delta_{i,(3-i)}$ corresponds to the transmission time in which a message is decoded at T_{3-i} (as in one-way case, a random variable with $E[\Delta_{i,(3-i)}] \leq N$.

Let an average rate-pair $(\bar{R}_{12}, \bar{R}_{21})$ be defined by the communication rates: $\bar{R}_{i,(3-i)} = \frac{\log|\mathcal{M}_i|}{\mathsf{E}[\Delta_{i,(3-i)}]}$ for $i = 1, 2$, and let the error probability in each direction be denoted as $\mathsf{P}_{e_{i,(3-i)}}\left(\bar{R}_{12}, \bar{R}_{21}, \Delta_{i,(3-i)}\right)$.

Definition 5: An error exponent pair $E_{i,(3-i)}\left(\bar{R}_{12}, \bar{R}_{21}\right)$ is achievable for a rate-pair $\left(\bar{R}_{12}, \bar{R}_{21}\right)$, over a two-way parallel DMC if there exists a sequence of two-way variable length codes such that $\mathsf{E}[\Delta_{i,(3-i)}] \leq N$ for $i = 1, 2$, and for very large N simultaneously:

$$\frac{-\log \mathsf{P}_{e_{i,(3-i)}}\left(\bar{R}_{12}, \bar{R}_{21}, \Delta_{i,(3-i)}\right)}{\mathsf{E}[\Delta_{i,(3-i)}]} \geq E_{i,(3-i)}\left(\bar{R}_{12}, \bar{R}_{21}\right).$$

Definition 6: The achievable error exponent region (EER) is the union over all achievable error exponent pairs at rate-pair $(\bar{R}_{12}, \bar{R}_{21})$.

III. Main Results: one-way

Consider a one-way DMC with $C_0 = 0$ and noiseless active feedback with rate-limited to R_{FB}[4]. Let any attainable error exponent for this channel at rate R in the absence of feedback be $E_{1w}(R)$. When noiseless feedback is used, improvements on the achievable error exponents depend on how R_{FB} compares to the forward expected rate $\bar{R}$, and how feedback is used to detect and correct errors. With Yamamoto-Itoh's [5] scheme, Burnashev's reliability is attained by feeding back the message decoding decision made at the receiver. However, the rate of the noiseless feedback transmission must equal that of the forward link only up to a critical rate $\bar{R}^*$, beyond which, as shown in [8], compressed noiseless feedback may be transmitted instead in the message mode of the Yamamoto-Itoh scheme in two forms: i) *random-hashing*: independently and uniformly assigning each of the messages into $2^{NR_{\text{FB}}}$ bins, whose index is fed back to the transmitter as a hash; and, ii) a *joint channel-code / hash-function* design where an erasure decoding rule takes into account the bins containing messages and is used to form a lower rate code. These approaches result in the following two propositions, as extensions of [8] that characterize achievable error exponents for a given noiseless feedback rate[5]:

[4]The noiseless feedback link has a capacity of $C_{0_{\text{FB}}}$, thus $R_{\text{FB}} \leq C_{0_{\text{FB}}}$.

[5]Error exponents are defined for the regime $R_\infty < \bar{R} \leq C$ if they depend on $E_{\text{sp}}(\cdot)$. In the regime $C_0 < \bar{R} \leq R_\infty$, $E_{1w}(R)$ is achievable without feedback using block codes. The reliability is unbounded for rates below C_0.

Proposition 1: An achievable error exponent for a one-way DMC with rate-limited R_{FB} noiseless active feedback, utilizing *random hashing* and VLC is given by $E(\bar{R}, R_{\text{FB}}) \geq E^{\text{RH}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}})$, for $\bar{R}_\infty \leq \bar{R} \leq C$, where:

$$E^{\text{RH}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}}) := \begin{cases} \max_{\bar{R} \leq R_{\text{data}} \leq C} \frac{\bar{R}}{R_{\text{data}}} E_{\text{Forn}}(R_{\text{data}}) & \\ \quad + \min\left\{ R_{\text{FB}}, \left(1 - \frac{\bar{R}}{R_{\text{data}}}\right) C_1 \right\}, & \text{if } R_{\text{FB}} < \bar{R}, \\ E_{\text{Burn}}(\bar{R}), & \text{if } R_{\text{FB}} \geq \bar{R}, \end{cases} \tag{3}$$

where $E_{\text{Forn}}(\cdot)$ corresponds to (2), and $E_{\text{Burn}}(\cdot)$ to (1). The maximization above applies for $\bar{R} < \bar{R}^*_c = \frac{CC_1}{C+C_1}$; when $\bar{R} \geq R^*_c$ then $R_{\text{data}} = C$.

Proof: See Appendix A in [10].

When the joint channel-coding / hashing-function method from [8, Sections 3.4-5] is used instead, we have the following:

Proposition 2: An achievable error exponent for a one-way DMC with rate-limited R_{FB} noiseless feedback, and *joint design channel-coding / hashing-function* and VLC is given by: $E(\bar{R}, R_{\text{FB}}) \geq E^{\text{Joint}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}})$ for $\bar{R}_\infty \leq \bar{R} \leq C$ as:

$$E^{\text{Joint}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}}) := \begin{cases} \min\left\{ R_{\text{FB}} + \frac{\bar{R}}{C} E_{\text{sp}}\left(\frac{C}{\bar{R}}(\bar{R} - R_{\text{FB}}), Q^*, W\right), \right. & \\ \quad \left. \left(1 - \frac{\bar{R}}{C}\right) C_1 \right\}, & \text{if } R_{\text{FB}} < \bar{R}, \\ E_{\text{Burn}}(\bar{R}), & \text{if } R_{\text{FB}} \geq \bar{R}, \end{cases} \tag{4}$$

where $E_{\text{sp}}(R, Q, W)$ corresponds to the sphere packing[6] error exponent for rate R, input distribution Q and channel law W, and Q^* is the capacity achieving input distribution.

Proof: Equation (4) results from [8, Equations (21-22)]. See Appendix B in [10].

Propositions 1 and 2 are based on the Yamamoto-Itoh scheme but using compressed noiseless feedback and allowing erasure decoding. Feedback is also exploited to maintain synchronization: the error-free control message informs the transmitter of whether the preliminary decision was accepted or not, regardless of correctness. The largest error exponent is attained by a hybrid system that chooses the scheme to use based on the rate-pair $(\bar{R}, R_{\text{FB}})$:

$$E_{\text{RL-FB}}\left(\bar{R}, R_{\text{FB}}\right) \geq \max\left\{ E^{\text{RH}}_{\text{RL-FB}}\left(\bar{R}, R_{\text{FB}}\right), E^{\text{Joint}}_{\text{RL-FB}}\left(\bar{R}, R_{\text{FB}}\right) \right\}.$$

Figure 1 shows this error exponent (vertical axis) for different values of forward and feedback rate-pairs (mapped on the horizontal plane). Note that for a fixed $\bar{R}$, as $0 \leq R_{\text{FB}} < \bar{R}$ the first line in the expressions of either Proposition 1 (green area) or 2 (gray area) are achievable. Once R_{FB} reaches $\bar{R}$, the reliability jumps to Burnashev's, which for low $\bar{R}$ occurs at an edge.

If the forward channel has $C_0 > 0$, the operation of the Yamamoto-Itoh scheme is simplified since the control stage is free of errors in both directions, thus we have:

Proposition 3: An achievable error exponent for a one-way DMC with rate-limited R_{FB} noiseless active feedback, using random hashing, VLC, and satisfying $0 < C_0 < \bar{R}$, is given by $E(\bar{R}, R_{\text{FB}}) \geq E^{\text{RH-C0}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}})$, where:

$$E^{\text{RH-C0}}_{\text{RL-FB}}(\bar{R}, R_{\text{FB}}) := \begin{cases} R_{\text{FB}} + E_{\text{Forn}}(\bar{R}), & \text{if } R_{\text{FB}} < \bar{R}, \\ & \text{for } \bar{R}_\infty \leq \bar{R} \leq C, \\ \infty, & \text{if } R_{\text{FB}} \geq \bar{R}, \\ & \text{for } 0 \leq \bar{R} \leq C. \end{cases} \tag{5}$$

Note that $E_{1w}(R)$ is achievable in the regime $C_0 < \bar{R} < R_\infty$, and recall that infinite reliability is attainable for $0 \leq \bar{R} \leq C_0$.

This proposition illustrates how $C_0 > 0$ in the forward channel may be exploited to boost reliability in the small error regime as a consequence of having perfect knowledge of the receiver's control mode decisions.

Proof: See Appendix C in [10].

For channels with noisy feedback, VLC strategies must use additional synchronization recovery techniques. In Yamamoto-Itoh like schemes, feedback control messages may be incorrectly decoded at the encoder, causing terminals to lose track of what message is being transmitted. In contrast, a single bit transmitted with zero-error in either direction suffices to maintain synchronization: i.e. a terminal may use this bit to signal the termination of its own message, or alternatively, that it accepted the current message sent from the other terminal.

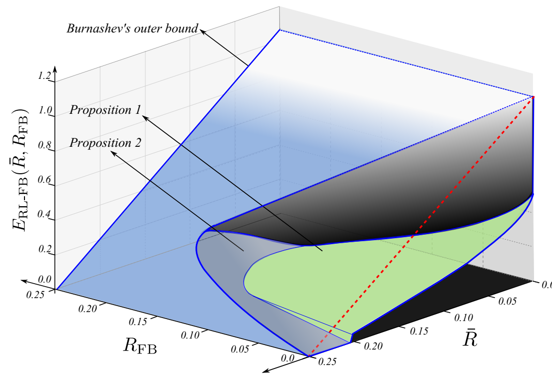

Fig. 1. Achievable Error exponent with rate-limited noiseless feedback and forward $C_0 = 0$. We evaluated all rate pairs satisfying $0 \leq \bar{R} \leq C$ and $0 \leq R_{\text{FB}} \leq C$ for a BSC(0.215).

IV. Main resuslts: two-way

1. Non-interacting terminals: Any error exponent achievable for a one-way DMC without feedback, $E_{1w}(R)$, is attainable in each direction of a two-way parallel DMC: i.e., $E_{12}(R_{12}, R_{21}) \geq E_{1w}(R_{12})$ and $E_{21}(R_{12}, R_{21}) \geq E_{1w}(R_{21})$.

2. Interacting terminals: when terminals employ feedback/ interaction, this affects the error exponents:

Proposition 4: An achievable error exponent pair for the two-way parallel DMC, in the rate-pair regime $0 < C_{0_{12}} < \bar{R}_{12} < C_{12}$ and $0 < C_{0_{21}} < \bar{R}_{21} < C_{21}$, using VLC is:

$$E_{12}(\bar{R}_{12}, \bar{R}_{21}) = E_{\text{Forn}}\left(\bar{R}_{12}\right)$$
$$E_{21}(\bar{R}_{12}, \bar{R}_{21}) = E_{\text{Forn}}\left(\bar{R}_{21}\right)$$

[6] $E_{\text{sp}}(R, Q, W)$ is used under the assumption of totally symmetric channels. See the discussion in [8, Equation (20), Sec. 3.4].

Proof: When $C_{0_{12}} = C_{0_{21}} = 0$, this is shown in [3, Prop. 3(i)]. Alternatively when $C_{0_{12}} > 0$ and $C_{0_{21}} > 0$, each terminal has access to at least a zero-rate noiseless feedback link, and Forney's (2) reliability can be directly achieved.

Next, we consider special cases where Proposition 4 can be further improved when noiseless feedback at positive rate is used in either direction depending on the zero-error capacity of each link. We first recall that Shannon [13] showed that in DMCs $R_\infty > 0$ if and only if every output cannot be reached from at least one channel input. Moreover, $0 \le C_0 \le R_\infty \le C$. Thus, depending on the transition probability matrix of a DMC, the following four cases (denoted by c_j for $j = 1, 2, 3, 4$) are possible in the two-way parallel DMC:

$$\begin{aligned} &c_1 : (C_0 = 0, R_\infty = 0), \quad c_2 : (C_0 = 0, R_\infty > 0), \\ &c_3 : (C_0 > 0, R_\infty > C_0), \quad c_4 : (C_0 = R_\infty > 0). \end{aligned} \tag{6}$$

There exist channels satisfying $R_\infty = C$, though we focus on the cases above. Thus, a two-way parallel DMC may result from the $1 \to 2$ link satisfying c_j, and the $1 \leftarrow 2$, c_k. We denote this by $(c_j; c_k)$ for $j, k \in \{1, \cdots, 4\}$. There are ten possible scenarios resulting from combinations of (6), two shown in Figure 2. Each may employ distinct schemes in different rate-pair regimes. We do not enumerate all possible cases; rather we present three examples that show how propositions of Section III can be used in the two-way setting.

In the first example, we show how the direction with $C_0 = 0$ benefits when the other has positive zero-error capacity. In the second and third examples, both channels have positive zero-error capacity. In the former the zero error capacities are equal and both directions benefit from it, exceeding Forney's reliability in subregion III. In the latter, zero-error capacities are different, yielding a small region outside rectangular regime $C_{0_{12}} \times C_{0_{21}}$ where infinite reliability is attainable.

Example 1. $(c_1; c_4)$: Consider that direction $1 \to 2$ is a symmetric channel with all probability entries strictly positive, (i.e. a row of the matrix is $[1-\epsilon, \frac{\epsilon}{2}, \frac{\epsilon}{2}]$ for $\epsilon > 0$) and $C_{0_{12}} = R_{\infty_{12}} = 0$. Direction $1 \leftarrow 2$ is a noisy typewriter channel [12] of 4 inputs with $C_{0_{21}} = R_{\infty_{21}} = 1$, and crossover probability $\epsilon < 1/2$. Thus $(c_1; c_4)$ results in the rate-pair regimes III and V as shown in Figure 2 (right). Regime V is further divided into two sub-regimes, V_I and V_{II} (see [10, Eq. (10)]), and a small portion labeled Ω. Region Ω is included in V_I or V_{II} depending on whether Proposition 1 or 2 is used for the $1 \to 2$ direction. For all rate-pairs in V, the $1 \leftarrow 2$ direction attains infinite reliability, thus we focus next on the $1 \to 2$ direction and illustrate how to take advantage of $C_{0_{21}} > 0$. We formalize this in the following:

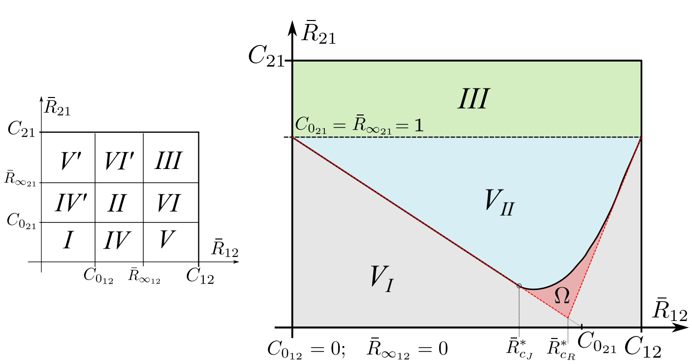

Fig. 2. Regimes in the capacity region of the two-way parallel DMC. Left: channels $(c_3; c_3)$, Right: channels $(c_1; c_4)$ –Example 1–.

Proposition 5: An achievable error exponent region for the two-way parallel DMC with $C_{0_{21}} = R_{\infty_{21}} > 0$ and $C_{0_{12}} = R_{\infty_{12}} = 0$ is determined for the following rate-pair regimes:

a. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in V_I$:

$$\begin{aligned} &E_{12}(\bar{R}_{12}, \bar{R}_{21}) = E_{\text{Burn}}(\bar{R}_{12}), \\ &E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty, \end{aligned}$$

b. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in V_{II}$:

$$\begin{aligned} &E_{12}(\bar{R}_{12}, \bar{R}_{21}) \ge E^{\text{Joint}}_{\text{RL-FB}}(\bar{R}_{12}, C_{0_{21}} - \bar{R}_{21}), \\ &E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty, \end{aligned}$$

when Proposition 2 is used, and if Proposition 1 is used:

$$\begin{aligned} &E_{12}(\bar{R}_{12}, \bar{R}_{21}) \ge E^{\text{RH}}_{\text{RL-FB}}(\bar{R}_{12}, C_{0_{21}} - \bar{R}_{21}), \\ &E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty, \end{aligned}$$

c. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in III$:

$$E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \begin{cases} \max_{\bar{R}_{12} \le R_{\text{data}_{12}} \le C_{12}} \frac{\bar{R}_{12}}{R_{\text{data}_{12}}} E_{\text{Forn}}(R_{\text{data}_{12}}) \\ \quad + \min\left\{ C_{0_{21}} \left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right), \left(1 - \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right) C_1 \right\}, \\ \qquad \text{if } C_{0_{21}} \left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right) < \bar{R}_{12}, \\ E_{\text{Burn}}(\bar{R}_{12}), \\ \qquad \text{if } C_{0_{21}} \left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right) \ge \bar{R}_{12}. \end{cases}$$

$$E_{21}\left(\bar{R}_{12}, \bar{R}_{21}\right) \ge \max_{\bar{R}_{21} \le R_{\text{data}_{21}} \le C_{21}} \frac{\bar{R}_{21}}{R_{\text{data}_{21}}} E_{\text{Forn}}\left(R_{\text{data}_{21}}\right).$$

Proof: See Appendix D in [10].

Example 2. $(c_4; c_4)$: Consider a two-way parallel DMC formed by two identical channels with positive zero error capacity. Each direction corresponds to a noisy typewriter channel of 4 inputs, with $C_{0_{12}} = R_{\infty_{12}} = C_{0_{21}} = R_{\infty_{21}} = 1$ and crossover probability $\epsilon < 1/2$. The resulting rate-regimes are subregions I, V, V' and III, from Figure 2 (left). Error exponent for regions V and V' follow similarly.

Proposition 6: An achievable error exponent region for the two way-parallel DMC with both directions having the same zero-error capacity is:

a. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in I$: $E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \infty, E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty$.
b. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in III$: See Proposition 7e in Example 3.
c. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in V$:

$$\begin{aligned} &E_{12}(\bar{R}_{12}, \bar{R}_{21}) \ge E_{\text{Forn}}(\bar{R}_{12}) + (C_{0_{21}} - \bar{R}_{21}), \\ &E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty. \end{aligned}$$

Proof: See Appendix E in [10]. Note that analogous results apply $\forall(\bar{R}_{12}, \bar{R}_{21}) \in V'$.

Example 3. $(c_3; c_4)$: Consider a two-way parallel DMC, both directions with positive zero-error capacity, one larger than the other. Let the $1 \to 2$ direction be a 4 input noisy typewriter channel with $C_{0_{12}} = R_\infty = 1$ and crossover probability $\epsilon < 1/2$; and let the $1 \leftarrow 2$ direction to be a pentagon channel [14] with $C_{0_{21}} = \log\sqrt{5} \approx 1.16$, and $R_\infty = \log\left(\frac{5}{2}\right)$. Figure 3 shows the capacity region and rate-pair regimes. The small triangle Γ achieves the same reliability as in I. Error exponents for V' result by flipping those of V, but Γ.

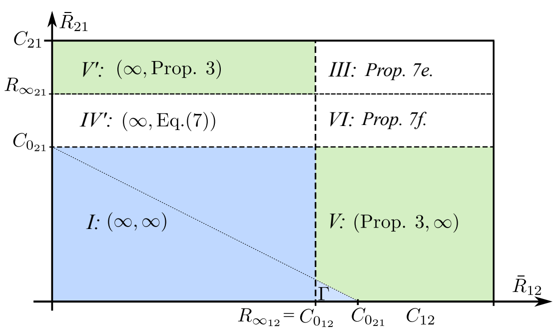

Fig. 3. Two-way parallel DMC of channel combination $(c_4; c_3)$.

Proposition 7: An achievable error exponent region for the two-way parallel DMC satisfying: $0 < C_{021} < R_{\infty_{21}} < C_{21}$, and $0 < C_{0_{12}} = R_{\infty_{12}} < C_{12}$ and $C_{0_{21}} > C_{0_{12}}$ is:

a. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in I$: $E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \infty, E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty$.

b. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in IV'$:

$$E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \infty$$
$$E_{21}(\bar{R}_{12}, \bar{R}_{21}) \geq E_r(\bar{R}_{21}) + (C_{0_{12}} - \bar{R}_{12}), \quad (7)$$

where (7) results from Proposition 3 with the random coding error exponent $E_r(R)$ instead of Forney's.

c. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in V$:

$$E_{12}(\bar{R}_{12}, \bar{R}_{21}) \geq E_{\text{Forn}}(\bar{R}_{12}) + (C_{0_{21}} - \bar{R}_{21}),$$
$$E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty.$$

d. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in \Gamma$: $E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \infty, E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \infty$.

e. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in III$: An achievable error exponent region results as Figure 4 (refer to Appendix F).

f. $\forall(\bar{R}_{12}, \bar{R}_{21}) \in VI$: This regime is essentially similar to III, with the distinction that the $1 \leftarrow 2$ direction cannot achieve Forney's reliability, thus, a general $E_{1w}(\cdot)$ error exponent should be used instead of $E_{\text{Forn}}(\cdot)$ since ML decoding is used.

Proof: See Appendix F in [10].

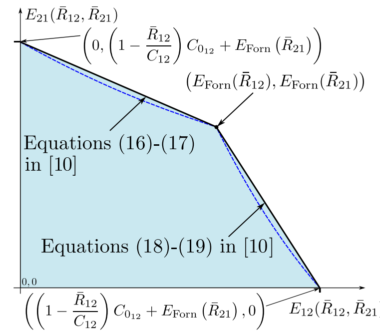

Fig. 4. Achievable EER for the subregion III of Proposition 7.

V. Discussion and Open Problems

There exist multiple open problems in two-way channels, including a) the two-way variable length zero-error capacity region; b) outer bounds for the two-way DMC EER for all rate-pair regimes; and c) how messages and feedback can be transmitted without invoking a time-sharing argument. Our initial characterization of EER aims to illustrate how a positive zero-error capacity can be exploited to not only resolve synchronization but also simplify coding schemes and increase reliability in the small error regime.

References

[1] C. E. Shannon, "Two-way communications channels," in *4th Berkeley Symp. Math. Stat. Prob.*, Chicago, IL, Jun. 1961, pp. 611–644.

[2] S. C. Draper and A. Sahai, "Variable-length channel coding with noisy feedback," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 355–370, Jun. 2008.

[3] K. Palacio-Baus, M. Asadi, and N. Devroye, "Error exponents of parallel two-way discrete memoryless channels using variable length coding," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2249–2253.

[4] M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," *Problemy Peredachi Informatsii*, vol. 12, no. 4, pp. 10–30, 1976.

[5] H. Yamamoto and K. Itoh, "Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback (Corresp.)," *IEEE Transactions on Information Theory*, vol. 25, no. 6, pp. 729–733, November 1979.

[6] G. Forney, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Transactions on Information Theory*, vol. 14, no. 2, pp. 206–220, Mar. 1968.

[7] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY: Wiley, 1968.

[8] S. C. Draper, K. Ramchandran, B. Rimoldi, A. Sahai, and D. Tse, "Attaining maximal reliability with minimal feedback via joint channel-code and hash-function design," *43th Annual Allerton conference on communication, control and computing*, 2005.

[9] A. Sato and H. Yamamoto, "Error exponents of discrete memoryless channels and AWGN channels with noisy feedback," in *2010 International Symposium On Information Theory and Its Applications*, Oct 2010, pp. 452–457.

[10] K. Palacio-Baus and N. Devroye, "Achievable error exponents for the two-way parallel DMC," https://devroye.lab.uic.edu/research-2/publications/, 2020, [Extended version].

[11] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 379-423, 623-656, Jul., Oct. 1948.

[12] ——, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, September 1956.

[13] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels - I," *Information and Control*, vol. 10, no. 1, pp. 65–103, 1967.

[14] L. Lovasz, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.

APPENDIX

A. *Proof of Proposition 1*

For the regime in which the feedback rate is smaller or equal to the forward expected rate, (3) is derived in a similar manner as [8, Equation (3)], by letting the receiver use erasure decoding instead, as we describe next. In the modified Yamamoto-Itoh scheme proposed in [8], the receiver feeds back the bin number (random hashing) that contains the estimated message using maximum likelihood[7]. When erasure decoding is used in addition to random hashing feedback, the receiver reserves a special message "$\{0\}$" (corresponding to bin index zero) to indicate the occurrence of an erasure and therefore, to request a message retransmission. If an erasure occurs, all subsequent transmissions (control mode in Yamamoto-Itoh) within this round are ignored by both terminals and the process starts over for the same message in the next round. If no erasure is declared by the end of the message mode, the scheme proceeds as in the modified Yamamoto-Itoh [8], with the difference that $2^{NR_{\text{FB}}}-1$ bins are available.

A block diagram of this coding scheme is shown in Figure 5. To fully use all channel uses in both directions, four streams of messages are interleaved. For each stream (i), a message chosen from the set $\{1,\cdots,2^{N\bar{R}}\}$ is transmitted in λN channel uses in the forward direction. Simultaneously, the hash message for message stream $i-1$ (corresponding to a hard decoding decision, either a bin number chosen from the set $\{1,\cdots,2^{NR_{\text{FB}}}-1\}$, or message $\{0\}$ for an erasure) is fed back in $N-\gamma$ channel uses, where $\gamma=\frac{1}{C_{0_{\text{FB}}}}$ does not depend on N. If a bin number is fed back, the encoder compares it with the one generated by the true message and replies a confirmation message (ACK/NACK), which is transmitted in the remaining $(1-\lambda)N$ channel uses (by means of a repetition code). This message (a single bit) is estimated by the receiver and returned to the encoder in γ channel uses without error.

We follow the notation of [8]. The probability of error is linked to the occurrence of an undetected error in the first stage of the message mode (p_u), and either the occurrence of a hash collision (p_h, given by the inverse of the number of bins) that is successfully acknowledged or a missed NACK (in the case there is no hash collision), $p_{n\to a}=2^{-N(1-\lambda)C_1}$ which is attained using a repetition code as in Yamamoto-Itoh [5] and by setting $p_{a\to n}\leq\delta$, where $\delta>0$ is an arbitrarily small number by Stein's lemma. Thus:

$$\begin{aligned}\mathsf{P}_e &= p_u\left[p_h p_{a\to a}+(1-p_h)p_{n\to a}\right]\\ &\leq p_u\left(p_h+p_{n\to a}\right)\\ &= 2^{-N\lambda E_{\text{Forn}}(R_{\text{data}})}\left[2^{-NR_{\text{FB}}}+2^{-N(1-\lambda)C_1}\right]\end{aligned}\tag{8}$$

where $\bar{R}\leq R_{\text{data}}\leq C$ corresponds to the instantaneous transmission rate in the forward direction and selected to maximize the attainable reliability; C_1 is determined as in (1). Above, we have upper bounded $p_u\leq 2^{-N\lambda E_{\text{Forn}}(R_{\text{data}})}$. Note that there are exactly $2^{(N-\gamma)R_{\text{FB}}}-1$ bins available, and that for large N and since γ is a fixed number, $2^{(N-\gamma)R_{\text{FB}}}-1\approx 2^{NR_{\text{FB}}}$.

Next, we analyze the probability of a retransmission P_{Rtx}, which occurs when the decoder declares either an erasure at the end of the message stage or when a NACK is declared at the end of the control mode. Thus, denoting the conditional probability that the preliminary decision is correct as $p_c=\mathsf{P}(M'=m|M=m)$, we have that:

$$\begin{aligned}\mathsf{P}_{\text{Rtx}} &= \mathsf{P}(\text{NACK or Erasure})\\ &\leq \mathsf{P}(\text{NACK})+\mathsf{P}(\text{Erasure})\\ &= p_c p_{a\to n}+p_u\left[p_h p_{a\to n}+(1-p_h)p_{n\to n}\right]+\mathsf{P}(\text{Erasure})\\ &\leq p_{a\to n}+p_u\left[p_h p_{a\to n}+p_{n\to n}\right]+\mathsf{P}(\text{Erasure})\\ &\leq p_{a\to n}+2p_u\leq\delta+2\left(2^{-\lambda N E_{\text{Forn}}(R_{\text{data}})}\right)+\mathsf{P}(\text{Erasure})\end{aligned}$$

Since δ can be chosen arbitrarily small, and since it can be shown that $\mathsf{P}(\text{Erasure})\to 0$ as $N\to\infty$ [6], the overall $\mathsf{P}_{\text{Rtx}}\to 0$ as $N\to\infty$. The probability of retransmission P_{Rtx} determines the number of block length N transmissions needed. This number is geometrically distributed with mean $\frac{1}{1-\mathsf{P}_{\text{Rtx}}}\approx 1$ when $\mathsf{P}_{\text{Rtx}}\to 0$. Hence, the expected transmission rate is $\bar{R}$, and the expected transmission time $\mathsf{E}[\Delta]=N$.

The error exponent resulting from (8) and the above is then:

$$\lim_{\mathsf{E}[\Delta]\to\infty}\frac{-\log_2\mathsf{P}_e}{\mathsf{E}[\Delta]}\geq\lambda E_{\text{Forn}}(R_{\text{data}})+\min\left\{R_{\text{FB}},(1-\lambda)C_1\right\}\tag{9}$$

Above, we pick $\lambda=\frac{\bar{R}}{R_{\text{data}}}$. Parameter R_{data} corresponds to the instantaneous transmission rate, which is chosen to maximize the error exponent as shown in Proposition 1. This choice in particular, affects the term on the left side of the sum in (9). The term on the right hand side of the sum is a function that increases linearly with R_{FB}, which dominates this term until R_{FB} matches the term $C_1\left(1-\frac{\bar{R}}{R_{\text{data}}}\right)$ (which in turn results in Burnashev's reliability when $R_{\text{data}}=C$). Next, note that the largest value that R_{FB} may take, before Burnashev's reliability is achievable, is to approach very closely, but not reaching $\bar{R}$. Thus, we can see the right term of (9) as a function that increases linearly with $\bar{R}$ instead, and until $\bar{R}=C_1\left(1-\frac{\bar{R}}{R_{\text{data}}}\right)$, which we denote as $\bar{R}=\bar{R}_c^*$. Beyond this point we set $R_{\text{data}}=C$ and the error exponent is totally characterized by the right term of the sum. For rates below $\bar{R}_c^*$, $R_{\text{data}}\neq C$ is chosen to maximize the error exponent (since in this regime $E_{\text{Forn}}(R_{\text{data}})\neq 0$).

The second line of (3) is achievable since with $R_{\text{FB}}\geq\bar{R}$, and the Yamamoto-Itoh scheme can be used directly.

B. *Proof of Proposition 2*

Similarly to Proposition 1, the error exponent for the regime in which the noiseless feedback rate equals or exceeds the

[7]Note that the schemes presented in [8] utilize bursty feedback, in which a single channel uses suffices to transmit a message from the receiver to the transmitter. Our schemes take into account that transmissions in the feedback direction occur in multiple channel uses, at a given rate that cannot exceed the feedback channel capacity.

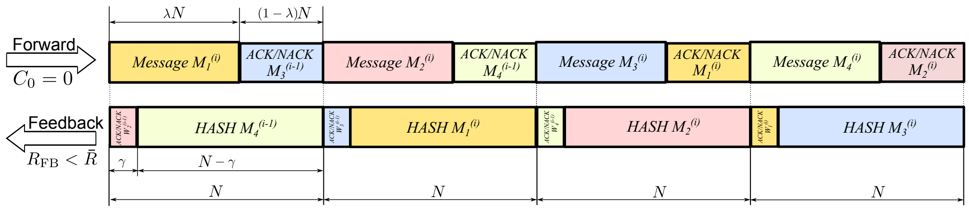

Fig. 5. One-Way DMC with rate-limited noiseless feedback: a block diagram of our coding scheme.

expected forward rate results by direct use of the Yamamoto-Itoh scheme. For the regime in which the feedback rate falls below the forward expected rate, (4) is obtained from the upper bound on the probability of error derived in [8, Sections 3.3-3.5]. This error exponent is achieved using the block diagram of Figure 5 and also used in achieving the error exponent in Proposition 1. Note that (4) is represented in terms of the sphere packing error exponent E_{sp}, as we have assumed totally symmetric channels.

C. Proof of Proposition 3

We use the Yamamoto-Itoh scheme, noticing that if $0 < C_0 < \bar{R}$, messages in the control mode may be exchanged without error. Hence, the receiver knows exactly whether the encoder agrees with the hash fed back in the message mode. Errors occur due to a hashing collision caused by a wrong preliminary decision (determined by erasure decoding and Forney's reliability), and that such incorrect decision results in the same bin as the true message (the probability of this event is given by the inverse of the number of bins).

If $R_{\text{FB}} \geq \bar{R}$, there are no hashing collision errors and the transmitter knows the decoder's preliminary decision (as in the original Yamamoto-Itoh scheme). Since there are no errors in the control mode, the true message can be retransmitted until it is received correctly, hence the second line of (5).

D. Proof of Proposition 5

Subregion III:

This rate-pair regime corresponds to both terminals transmitting at rates beyond the one-way zero-error capacity of their corresponding forward channels. In this example $C_{0_{21}} > 0$ and $C_{0_{12}} = 0$. Proposition 5 corresponds to the use of Proposition 1 for the $1 \to 2$ direction and the one-way scheme with noisy feedback from [2] for the other ($1 \leftarrow 2$) direction. The scheme operates as shown in the block diagram of Figure 6 which is a modification of the Yamamoto-Itoh coding scheme that allows two-way message transmission. Note that message interleaving has been used in order to make better use of both directions.

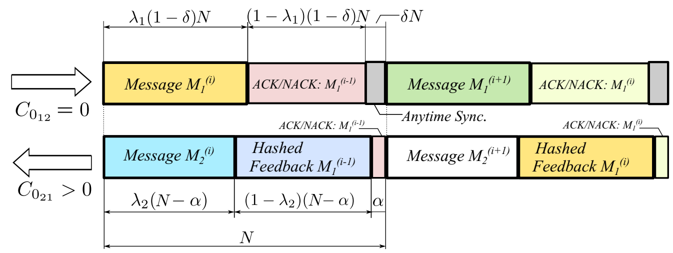

Fig. 6. Block Diagram for subregion III of Proposition 5.

For the $1 \to 2$ direction, the use of Proposition 1 occurs as follows: a message M_1 is sent in $\lambda_1(1-\delta)N$ channel uses, then a feedback hash message is returned in the opposite direction with zero-error in $(1-\lambda_2)(N-\alpha)$ channel uses and rate equal to $C_{0_{21}}$. Depending on whether the received hash matches the one generated by the true message, terminal 1 responds with a ACK/NACK repetition code in $(1-\lambda_1)(1-\delta)N$ channel uses. This message is decoded by terminal 2 and later fed back with no error in $\alpha = 1/C_{0_{21}}$ channel uses. The protocol operates seamlessly as for one-way.

For the $1 \leftarrow 2$ direction Forney's reliability can be attained by means of the coding scheme proposed in [2]. This scheme makes use of a round robin message scheduling that transmits messages from L data stacks by means of an anytime synchronization code of zero rate. This technique allows conveying a single bit over a noisy channel with reliability that increases with L. Note that terminal 2 transmit message M_2 in $\lambda_2(N-\alpha)$ channel uses. The anytime code is sent in the $1 \to 2$ direction using δN channel uses. Here, $\delta > 0$ corresponds to an arbitrarily small positive number such that $\lim_{N\to\infty} \frac{1}{\delta N} = 0$.

Note that the coding scheme must ensure first that a target average rate-pair $(\bar{R}_{12}, \bar{R}_{21})$ is achieved. To this end, parameters λ_1 and λ_2 must be chosen so that each direction communicates at the desired rate. We choose these two parameters as:

$$\lambda_1 = \frac{\bar{R}_{12}}{R_{\text{data}_{12}}} \quad \text{and} \quad \lambda_2 = \frac{\bar{R}_{21}}{R_{\text{data}_{21}}},$$

where $\bar{R}_{12} \leq R_{\text{data}_{12}} \leq C_{12}$ and $\bar{R}_{12} \leq R_{\text{data}_{21}} \leq C_{21}$ correspond to the effective instantaneous rate in each direction respectively, and selected to attain the largest error exponent determined by the scheme.

Then, from (3) an achievable error exponent for the $1 \to 2$ direction is:

$$E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \begin{cases} \max_{\bar{R}_{12} \le R_{\text{data}_{12}} \le C_{12}} \frac{\bar{R}_{12}}{R_{\text{data}_{12}}} E_{\text{Forn}}(R_{\text{data}_{12}}) \\ \quad + \min\left\{ C_{0_{21}}\left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right), \left(1 - \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right) C_1 \right\}, \\ \qquad\qquad \text{if } C_{0_{21}}\left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right) < \bar{R}_{12}, \\ E_{\text{Burn}}(\bar{R}), \\ \qquad\qquad \text{if } C_{0_{21}}\left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right) \ge \bar{R}_{12}. \end{cases}$$

whereas, an achievable error exponent for the $1 \leftarrow 2$ direction is given by:

$$E_{21}\left(\bar{R}_{12}, \bar{R}_{21}\right) \ge \max_{\bar{R}_{21} \le R_{\text{data}_{21}} \le C_{21}} \frac{\bar{R}_{21}}{R_{\text{data}_{21}}} E_{\text{Forn}}\left(R_{\text{data}_{21}}\right).$$

Subregion V: The achievable error exponents in V_I and V_{II} result from direct application of Propositions 1 and 2. To understand how subregion V is partitioned consider the error exponents in the $1 \to 2$ direction: when Proposition 2 is used, V_I corresponds to all rate-pairs $(\bar{R}_{12}, \bar{R}_{21})$ satisfying:

$$\bar{R}_{21} \le C_{0_{21}} - \bar{R}^*_{\text{FB}}, \tag{10}$$

where, from [8, Equation (22)], $\bar{R}^*_{\text{FB}}$ corresponds to:

$$\bar{R}^*_{\text{FB}} = \min\{\bar{R}_{12}, r\},$$

and $r > 0$ is the smallest positive rate that satisfies:

$$r + \frac{\bar{R}_{12}}{C_{12}} E_{\text{sp}}\left(\frac{C_{12}}{\bar{R}_{12}}(\bar{R}_{12} - r), Q^*, W_{12}\right) \ge C_1\left(1 - \frac{\bar{R}_{12}}{C_{12}}\right),$$

for all $0 \le \bar{R}_{12} \le C_{12}$. In this characterization the small portion Ω is part of V_I. Alternatively, when Proposition 1 is used, V_I includes all rate-pairs $(\bar{R}_{12}, \bar{R}_{21})$ satisfying (10) with:

$$\bar{R}^*_{\text{FB}} = \min\left\{\bar{R}_{12}, C_1\left(1 - \frac{\bar{R}_{12}}{C_{12}}\right)\right\}$$

for all $0 \le \bar{R}_{12} \le C_{12}$. Now, region Ω is part of V_{II}.

The corresponding error exponents result as:

- V_I: The $1 \to 2$ direction can achieve Burnashev's bound, since for every rate-pair in F_I, T_2 can provide noiseless decision feedback to terminal T_1.
- V_{II}: The error exponent in the $1 \to 2$ direction results from Propositions 1 and 2, since for both, T_2 can provide noiseless feedback to T_1 at a rate of $C_{0_{21}} - \bar{R}_{21}$.

The point $\bar{R} = R^*_{c,J}$ in Figure 2 indicates the largest rate for which both methods require the same feedback rate to attain Burnashev's bound in the $1 \to 2$ direction. Beyond that point, Proposition 1 requires a larger feedback rate than in Proposition 2.

We now present a coding scheme for region V, based on the block diagram of Figure 7, which modifies Yamamoto-Itoh's scheme to support two-way communication. First, message transmissions for both terminals have been interleaved to make more efficient use of both channels. For V_I, the scheme is used to achieve Burnashev's bound in the $1 \to 2$ direction. Observe that with $\alpha = 1/C_{0_{21}}$ channel uses it is possible to convey ACK/NACK confirmation messages for the feedback control mode with zero-error in the $1 \leftarrow 2$ direction. The message transmission and feedback stages for the $1 \to 2$ direction occur in $\lambda(N-\alpha)$ channel uses. Message transmission for the $1 \leftarrow 2$ direction occurs in $(1-\lambda)(N-\alpha)$ channel uses. This means that λ is the time sharing parameter, which controls the number of channel uses available for messages transmission in the noiseless channel ($1 \leftarrow 2$). This channel transmits with zero-error at an instantaneous rate of $C_{0_{21}}$. Once the target average rate $\bar{R}_{21}$ is established, one can determine the value of λ necessary to obtain such rate as $\bar{R}_{21} = \lambda C_{0_{21}}$. This leaves a rate of $(1-\lambda)C_{0_{21}}$ to provide feedback for the other direction, alternatively represented as $C_{0_{21}} - \bar{R}_{21}$. The operation of the scheme remains unchanged regardless of whether decision or hashed feedback is sent.

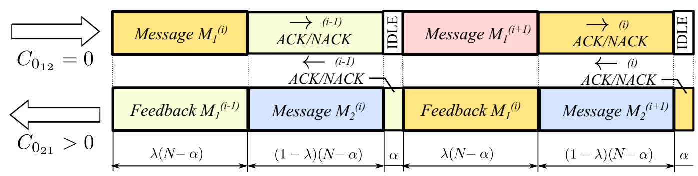

Fig. 7. Block diagram for the two-way parallel DMC with one direction having strictly positive zero error capacity.

E. Proof of Proposition 6

The results of Proposition 6 can be achieved by means of the block diagram of Figure 8:

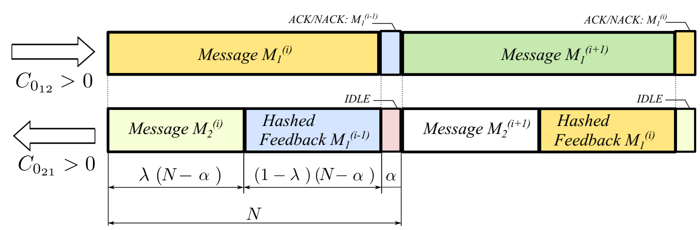

Fig. 8. Block Diagram for subregion V

In the $1 \to 2$ direction, message M_1 is transmitted in $(N-\alpha)$ channel uses, at a rate $\bar{R}_{12}$. In the other direction, message M_2 is transmitted without error in $\lambda(N-\alpha)$ channel uses. Here we choose $\lambda = \frac{\bar{R}_{21}}{C_{0_{21}}}$, such that a target average rate $\bar{R}_{21}$ can be attained transmitting at an instantaneous rate equal to $C_{0_{21}}$ when the block length is shortened by λ. In the remaining $(1-\lambda)(N-\alpha)$ channel uses, a hash regarding the message sent by T_1 is fed back without error at a rate of $C_{0_{21}} - \bar{R}_{21}$. Since we use the same principles as in the Yamamoto-Itoh scheme, a confirmation message needs to be sent in the control mode to T_2. Exploiting the zero-error capacity $C_{0_{12}}$, a codeword of length $\alpha = \frac{1}{C_{0_{12}}}$ indicates whether a retransmission is necessary or not. This message needs no confirmation as it is sent without error.

F. Proof of Proposition 7

Subregion Γ: In the $1 \leftarrow 2$ direction, all rates are below $C_{0_{21}}$ whereas for the other direction infinite reliability is attained even for rates beyond $C_{0_{12}}$. This is possible since perfect decision feedback may be provided to T_1 for all rates satisfying $C_{0_{12}} \leq \bar{R}_{12} \leq C_{0_{21}}$, since $C_{0_{21}} > C_{0_{12}}$. Thus, following Yamamoto-Itoh's scheme, and since there are no errors in the control mode messages (both directions have positive zero error capacity), the message in the $1 \rightarrow 2$ direction can be retransmitted until it is received perfectly. It can be shown that the number of retransmission has a geometric distribution and that the expected number of transmissions is one.

Subregion III: Forney's reliability is achievable by taking advantage of error-free zero-rate transmissions that are possible when both directions have positive zero-error capacity. These are used to send retransmission requests resulting from an erasure output in a decoder.

The above may be further improved by exploiting the zero-error capacity of each direction at positive rate. We consider a coding scheme operating according to the following block diagram similar to the Yamamoto-Itoh scheme:

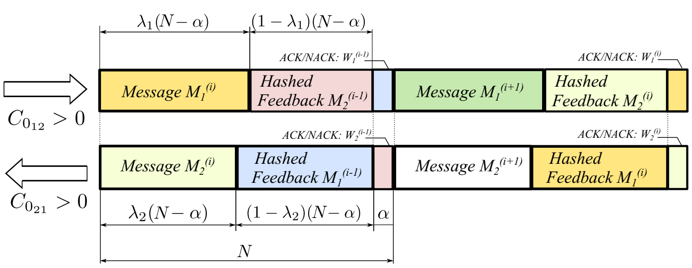

Fig. 9. Block Diagram for subregion V

The scheme comprises three stages: i) *message transmission*, in which messages M_i are transmitted in $\lambda_i(N-\alpha_i)$ channel uses using a capacity achieving code, and estimated at each receiver respectively; ii) *hashed feedback*, lasting for $(1-\lambda_i)(N-\alpha_i)$ channel uses, in which a hash derived from the estimated message is fed back without error; and iii) *confirmation* (for the control mode message), which lasts $\alpha_i = \frac{1}{C_{0_{i,(3-i)}}}$ channel uses, in which a single bit is sent from terminal T_i to T_{3-i} to indicate whether the preliminary decision hash is correct. If not, a retransmission follows. Note that since control messages are transmitted without error, no acknowledgment of this bit is necessary in the opposite direction.

Assume that both channels operate at the average rate-pair $(\bar{R}_{12}, \bar{R}_{21})$. Next, consider the $1 \rightarrow 2$ direction (results for the other direction follow by symmetry). With random hashing, each of the $2^{(N-\alpha_1)\bar{R}_{12}}$ messages are independently and uniformly assigned to $2^{(1-\lambda_2)(N-\alpha_2)C_{0_{21}}}$ bins. The number of bins is determined by the amount of channel uses that remain once message M_2 has been transmitted in $\lambda_2(N-\alpha_2)$ channel uses, and the $1 \leftarrow 2$ channel's zero-error capacity (since it determines the largest number of messages that can be fed back). Parameter λ_2 is determined by the average rate $\bar{R}_{21}$ at which the $1 \leftarrow 2$ direction operates. Since the block length is shortened by λ_2, a target $\bar{R}_{21}$ is reached by transmitting at a rate $\bar{R}_{21} \leq R_{\text{data}_{21}} < C_{21}$. Then, since for very large N we have that $N - \alpha_2 \approx n$, we can set $\lambda_2 = \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}$.

The hash is used once the receiver makes a preliminary decision. The corresponding hash (bin number) is fed back without error, so the transmitter can determine if a retransmission is necessary when the received hash does not match the one of the true message. We let receivers use erasure decoding and reserve message (bin) $\{0\}$ to indicate an erasure, thus enabling decoders to initiate a retransmission.

An analogous analysis follows when hashing in the other direction. Note that the block diagram of Figure 9 depicts the transmission of two messages of the data stream which have been interleaved to make better use of the channels in both directions. Assume $M_i = m$ is sent, then, an overall error occurs given the event that the preliminary decision M_i' is not correct and the bin number to which M_i' is assigned is the same as the one of the true message M_i. Let $h(m)$ denote the hash generated by message m, then the probability of error is:

$$\begin{aligned}
\mathsf{P}_{e_{12}} &= \mathsf{P}\left(M_1' = m', h(m') = h(m) \mid M_1 = m\right) \\
&= \mathsf{P}\left(h(m') = h(m) \mid M_1' = m', M_1 = m\right) \\
&\quad \cdot \mathsf{P}\left(M_1' = m' \mid M_1 = m\right) \qquad (11)
\end{aligned}$$

In (11), the first factor corresponds to the probability of a hash collision, whereas the second to a wrong decoding during the transmission stage (an incorrect preliminary decision) which is determined by Forney's reliability. Thus, we can upper bound this probability as:

$$\mathsf{P}_{e_{12}} \leq 2^{-(N-\alpha_2)(1-\lambda_2)C_{0_{21}}} \cdot 2^{-\lambda_1 N E_{\text{Forn}}(R_{\text{data}_{12}})}. \qquad (12)$$

Equivalently, for the other direction:

$$\mathsf{P}_{e_{21}} \leq 2^{-(N-\alpha_1)(1-\lambda_1)C_{0_{12}}} \cdot 2^{-\lambda_2 N E_{\text{Forn}}(R_{\text{data}_{21}})}. \qquad (13)$$

Since $\lambda_1 = \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}$ and $\lambda_2 = \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}$, (12) and (13) become:

$$\mathsf{P}_{e_{12}} \leq 2^{-(N-\alpha_2)\left(1-\frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right)C_{0_{21}}} \cdot 2^{-\frac{\bar{R}_{12}}{R_{\text{data}_{12}}} N E_{\text{Forn}}(R_{\text{data}_{12}})},$$

$$\mathsf{P}_{e_{21}} \leq 2^{-(N-\alpha_1)\left(1-\frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right)C_{0_{12}}} \cdot 2^{-\frac{\bar{R}_{21}}{R_{\text{data}_{21}}} N E_{\text{Forn}}(R_{\text{data}_{21}})},$$

for $\bar{R}_{12} \leq R_{\text{data}_{12}} \leq C_{12}$ and $\bar{R}_{21} \leq R_{\text{data}_{21}} \leq C_{21}$.

It can be shown that the expected transmission time $\mathsf{E}[\Delta_{12}] = \mathsf{E}[\Delta_{21}] = N$. Then, the corresponding error exponents result from:

$$\lim_{\mathsf{E}[\Delta_{12}]\to\infty} \frac{-\log_2 \mathsf{P}_{e_{12}}}{\mathsf{E}[\Delta_{12}]} \geq E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \left(1 - \frac{\bar{R}_{21}}{R_{\text{data}_{21}}}\right) C_{0_{21}} + \frac{\bar{R}_{12}}{R_{\text{data}_{12}}} E_{\text{Forn}}\left(R_{\text{data}_{12}}\right) \qquad (14)$$

$$\lim_{\mathsf{E}[\Delta_{21}]\to\infty} \frac{-\log_2 P_{e_{21}}}{\mathsf{E}[\Delta_{21}]} \geq E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \left(1 - \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right) C_{0_{12}} + \frac{\bar{R}_{21}}{R_{\text{data}_{21}}} E_{\text{Forn}}\left(R_{\text{data}_{21}}\right) \qquad (15)$$

Equations (14) and (15) show that there exists a trade off for the choice of the instantaneous rate $R_{\text{data}_{i,(3-i)}}$ in each direction. The instantaneous rate is selected first to guarantee that an expected communication rate is attained, and second, to benefit one's own or the other direction error exponent. Observe that the error exponent in the $1 \to 2$ direction, if $R_{\text{data}_{12}} = \bar{R}_{12}$ the term $\frac{\bar{R}_{12}}{R_{\text{data}_{12}}} E_{\text{Forn}}\left(R_{\text{data}_{12}}\right)$ in (14) is maximized. However, this causes that $\left(1 - \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right) C_{0_{12}} = 0$ in (15), reducing the achievable error exponent in the $1 \leftarrow 2$ direction. Alternatively, the choice $R_{\text{data}_{12}} = C_{12}$ benefits the $1 \leftarrow 2$ direction since it maximizes the term $\left(1 - \frac{\bar{R}_{12}}{R_{\text{data}_{12}}}\right) C_{0_{12}}$ in (15), so that the achievable error exponent $E_{12}(\bar{R}_{12}, \bar{R}_{21})$ depends only on the choice of $R_{\text{data}_{21}}$, since the term $\frac{\bar{R}_{12}}{R_{\text{data}_{12}}} E_{\text{Forn}}\left(R_{\text{data}_{12}}\right) = 0$ in (14). These two possible choices for $R_{\text{data}_{i,(3-i)}}$, can be seen as two different coding schemes in which: (i) Forney's reliability can be achieved in both directions, and (ii) both terminals cooperate to maximize the error exponent of a specific direction. We describe them as follows: choosing $R_{\text{data}_{12}} = \bar{R}_{12}$ and $R_{\text{data}_{21}} = \bar{R}_{21}$ leads to (i), which is indicated using a superscript:

$$
\begin{aligned}
E_{12}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) &= E_{\text{Forn}}\left(\bar{R}_{12}\right) \\
E_{21}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) &= E_{\text{Forn}}\left(\bar{R}_{21}\right)
\end{aligned}
$$

There exist two cases for (ii), one in which the error exponent in the $1 \to 2$ direction is maximized, and another in which the $1 \leftarrow 2$ direction is maximized. We denote these choices correspondingly by (ii_{12}) and (ii_{21}). Choosing $R_{\text{data}_{12}} = C_{12}$ and $R_{\text{data}_{21}} = \bar{R}_{21}$, the achievable error exponent region is:

$$
\begin{aligned}
E_{12}^{(ii_{12})}(\bar{R}_{12}, \bar{R}_{21}) &= 0 \\
E_{21}^{(ii_{12})}(\bar{R}_{12}, \bar{R}_{21}) &= \left(1 - \frac{\bar{R}_{12}}{C_{12}}\right) C_{0_{12}} + E_{\text{Forn}}\left(\bar{R}_{21}\right),
\end{aligned}
$$

equivalently in (ii_{21}), $R_{\text{data}_{12}} = \bar{R}_{12}$ and $R_{\text{data}_{21}} = C_{21}$:

$$
\begin{aligned}
E_{12}^{(ii_{21})}(\bar{R}_{12}, \bar{R}_{21}) &= \left(1 - \frac{\bar{R}_{21}}{C_{21}}\right) C_{0_{21}} + E_{\text{Forn}}\left(\bar{R}_{12}\right) \\
E_{21}^{(ii_{21})}(\bar{R}_{12}, \bar{R}_{21}) &= 0.
\end{aligned}
$$

Time sharing between these schemes (i) and (ii) achieves the error exponent region characterized below and shown in solid black line in Figure 4 (dashed blue result from (14) and (15) for all possible choices of $R_{\text{data}_{i,(3-i)}}$). When schemes (i) and (ii_{12}) are time-shared by parameter $0 \le \alpha \le 1$ we have (16) and (17):

$$
E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \alpha E_{12}^{(ii_{12})}(\bar{R}_{12}, \bar{R}_{21}) + (1-\alpha) E_{12}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) \tag{16}
$$

$$
E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \alpha E_{21}^{(ii_{12})}(\bar{R}_{12}, \bar{R}_{21}) + (1-\alpha) E_{21}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) \tag{17}
$$

Alternatively for (i) and (ii_{21}) and parameter $0 \le \beta \le 1$ we have (18) and (19):

$$
E_{12}(\bar{R}_{12}, \bar{R}_{21}) = \beta E_{12}^{(ii_{21})}(\bar{R}_{12}, \bar{R}_{21}) + (1-\beta) E_{12}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) \tag{18}
$$

$$
E_{21}(\bar{R}_{12}, \bar{R}_{21}) = \beta E_{21}^{(ii_{21})}(\bar{R}_{12}, \bar{R}_{21}) + (1-\beta) E_{21}^{(i)}(\bar{R}_{12}, \bar{R}_{21}) \tag{19}
$$

Observe that both schemes can be multiplexed without affecting the expected transmission rate $\bar{R}$ in both directions, since R_{data} is selected in each scheme to guarantee the same target average rate is achieved.