Abstract—We present a new technique to obtain outer-bounds on the capacity region of networks with ultra low-rate feedback. We establish a connection between the achievable rates in the forward channel and the minimum distortion that can be attained over the feedback channel.

Index Terms—Rate-limited feedback, erasure channels, channel capacity, Shannon feedback, sub-bit feedback.

I. INTRODUCTION

The introduction of massive Machine-Type Communications (mMTC) challenges many assumptions taken for granted in network information theory. One of the main challenges is the increased cost of learning. For instance, in current systems, it is well justified to assume free access to small control packets (e.g., ACK/NACK signals) when needed, as they are much smaller in size compared to payload packets. However, this is no longer the case in mMTC where payload and control packets will be of comparable sizes. To make matters more complicated, there is a growing concern about security attacks that aim to disrupt unprotected feedback channels.

We present new outer-bounds for networks with low-rate feedback. The outer-bound establishes a connection between rate-distortion theory and achievable rates in multi-terminal networks with feedback. Interestingly, we learn that the best use of the feedback channel may not be to minimize the error but rather the distortion in reconstructing channel information. More specifically, we first quantify how closely the channel state may be reconstructed using the rate-limited feedback link, and then, we define a space of indistinguishable channel realizations in which the outer-bound is optimized.

To illustrate the technique, we focus on the two-user broadcast erasure channel (BEC) and assume a sub-bit feedback link from only one receiver, $\text{Rx}_F$, while the other receiver, $\text{Rx}_N$, does not share any information with the other nodes. $\text{Rx}_F$ uses the rate-limited feedback link to causally provide its potentially encoded channel state information (CSI) to the other nodes. Although it remains open whether the bound is tight, it is the first of its kind and the surprising message is that to achieve the bound, the approach may be to purposefully induce distortion into the encoded CSI. We outline how existing results fall short of achieving these bounds and provide further insights and interpretations.

Related Literature: In [1], outer-bounds for wiretap channels with rate-limited output feedback were derived, which are tight for physically degraded channels. Shayevitz and Wigger observed in [2] that finding a general feedback capacity formula for memoryless broadcast channels (BCs) is very hard. For other channels such as Gaussian BCs, the capacity with single-user feedback is still unknown [3]. The block Markov coding of [4] provides interesting insights but involves characterizing complicated auxiliary random variables. To further understand how low-rate feedback affects the capacity region of multi-terminal channels, two-user broadcast erasure channels with intermittent [5], [6] and one-sided [7]–[9] delayed feedback have been studied. Interestingly, it was shown in [9] that even when only one receiver provides its delayed CSI to the transmitter, the outer-bound with global delayed feedback

A. Vahid is with the Department of Electrical Engineering, University of Colorado Denver, USA. Email: alireza.vahid@ucdenver.edu.
can be achieved. This latter finding motivates us to further lower the feedback rate to sub-bit territory to understand the fundamental limits of communications with rate-limited feedback in its purest form. In the context of interference channels, in [11], [12], the capacity of two-user interference channels with rate-limited feedback was established; new coding schemes with noisy or intermittent output feedback were proposed in [13], [14]; and [15]–[17] generalize such ideas to two-way communications. Finally, locality of feedback was studied under different delay assumptions in [18]–[24].

II. PROBLEM FORMULATION

We consider the two-user broadcast erasure channel (BEC) of Fig. 1 in which a single-antenna transmitter, Tx, wishes to communicate two independent messages, \( W_F \) and \( W_N \), to two single-antenna receiving terminals \( \text{Rx}_F \) and \( \text{Rx}_N \) (read feedback/silent receiver), respectively, over \( n \) channel uses. Each of the messages, \( W_F \) and \( W_N \), is uniformly distributed over \( \{1, 2, \ldots, 2^{nR_F}\} \) and \( \{1, 2, \ldots, 2^{nR_N}\} \), respectively. At time instant \( t \), the messages are mapped to channel input \( X[t] \in \mathbb{F}_2 \) (in the binary field), and the respective received signals at \( \text{Rx}_F \) and \( \text{Rx}_N \) are:

\[
Y_F[t] = S_F[t]X[t] \quad \text{and} \quad Y_N[t] = S_N[t]X[t],
\]

where \( \{S_F[t]\} \) is the Bernoulli \( (1 - \delta_F) \) process that governs the erasure at \( \text{Rx}_F \), and \( \{S_N[t]\} \) is the Bernoulli \( (1 - \delta_N) \) process that governs the erasure at \( \text{Rx}_N \). In this manuscript, we assume the channels are distributed independently over time and across users, and we limit the scope to \( \delta_F = \delta_N = \delta \) (i.e., homogeneous channels) where \( \delta \) is known globally.

We assume that at time instant \( t \), each receiver knows its channel value instantly, e.g., at time instant \( t \), \( \text{Rx}_F \) knows the realization of \( S_F[t] \). When the channel realization is 1, the corresponding terminal receives \( X[t] \) noiselessly, and when it is 0, the terminal understands an erasure has occurred.

We further assume a one-sided feedback structure in which at time instant \( t \), \( \text{Rx}_N \) knows \( S_N[t] \) but does not share this information with the other nodes. On the other hand, we assume \( \text{Rx}_F \) shares its channel state with receiver \( \text{Rx}_N \) and the transmitter through a rate-limited feedback channel of capacity \( C_{FB} \). We assume the feedback encoder at \( \text{Rx}_F \) sends back feedback symbol \( K[t] \in \mathcal{K}[t] \) at time \( t \) and this information becomes available to the other nodes at time \( t + 1 \). Here, \( \mathcal{K}[t] \) is the feedback alphabet at time \( t \). The feedback symbol depends causally on \( S_F[t] \), and the cardinality of the feedback alphabets satisfies

\[
\frac{1}{t} \sum_{\ell=1}^{t} \log_2 (|\mathcal{K}[\ell]|) \leq C_{FB}, \quad \forall 1 \leq t \leq n.
\]

The constraint imposed at the encoding function \( f_\ell(.) \) at time index \( \ell \) is

\[
X[t] = f_\ell (W_F, W_N, K^\ell-1).
\]

Receivers \( \text{Rx}_F \) and \( \text{Rx}_N \) use decoding functions \( \varphi_{F,n} (Y_F^n, S_F^n) \) and \( \varphi_{N,n} (Y_N^n, K^n, S_N^n) \) to get estimates \( \hat{W}_F \) of \( W_F \) and \( \hat{W}_N \) of \( W_N \), respectively. An error occurs whenever the estimate does not match the corresponding message. The average probabilities of error are given by

\[
\lambda_{F,n} = \mathbb{E}[P(\hat{W}_F \neq W_F)], \quad \lambda_{N,n} = \mathbb{E}[P(\hat{W}_N \neq W_N)],
\]

where the expectations are taken with respect to the random choice of the transmitted messages.

We say that a rate pair \((R_F, R_N)\) is achievable if there exists a block encoder at the transmitter, and a block decoder at each receiver, such that the average probabilities of error go to zero as the block length \( n \) goes to infinity. The capacity region, \( \mathcal{C} \), is the closure of the set of achievable rate pairs.

![Fig. 1. Two-user BEC with rate-limited one-sided feedback.](image-url)
III. MAIN RESULTS

In this section, we provide the new distortion-based outer-bound for the two-user BEC with rate-limited feedback. Define \( 0 \leq D^* \leq \min \{ \delta, 1 - \delta \} \) to be the unique value to satisfy

\[
H(D) = (H(\delta) - C_{\text{FB}})^+,
\]

and \( \gamma_{\text{out}} \) to be

\[
\gamma_{\text{out}} \triangleq \frac{D^*}{\min \{ \delta, 1 - \delta \}}.
\]

**Theorem 1.** The capacity region, \( \mathcal{C} \), of the two-user BEC with one-sided rate-limited feedback is included in

\[
\mathcal{C}_{\text{out}} \triangleq \left\{ (R_F, R_N) \left| \begin{array}{c}
R_F + \beta_{\text{out}} R_N \leq \gamma_{\text{out}} (1 - \delta) \\
\beta_{\text{out}} R_F + R_N \leq \gamma_{\text{out}} (1 + \delta)
\end{array} \right. \right\},
\]

where

\[
\beta_{\text{out}} = \gamma_{\text{out}} + (1 - \gamma_{\text{out}}) (1 + \delta).
\]

Similar to the findings of [9], although only \( \text{Rx}_F \) provides feedback, the outer-bound is symmetric. The results could also provide new capacity bounds for erasure interference channels based on their connection to BECs [25]–[28].

Figure 2 plots the outer-bound on the maximum sum-rate point for \( \delta = 0.4 \). When \( C_{\text{FB}} \geq H(\delta) \), this outer-bound matches the capacity region of the two-user BEC with global delayed feedback [29]. This recovers the results of [9], [10] where it is shown that perfect one-sided feedback is as good as global feedback. The interesting distinction between the results in [9], [10] and prior results is the fact that delayed CSI is harnessed at every step of the achievability. One can envision a Markov block structure to interleave different blocks and compress the feedback to \( H(\delta)n \) bits and mimic the results of [9] to achieve of point A in Figure 2 where \( H(\delta) = C_{\text{FB}} \). However, below this limit, the achievability can no longer be derived from such arguments. In fact, the argument presented above requires perfect delayed CSI, and for \( C_{\text{FB}} < H(\delta) \), if we attempt to send perfectly a part of CSI back to the transmitter, we deviate from this outer-bound as in Figure 2.

This suggests the possibility of new achievability ideas that purposefully induce distortion in the encoded CSI.

![Fig. 2. Sum-rate outer-bound for \( \delta = 0.4 \). Point A is where \( H(\delta) = C_{\text{FB}} \).](image-url)
Proof. We divide the proof into two parts based on the conditions expressed in Claim 1.

Step 2: In this step, we claim if we replace sequence $S_F^n$ with $\hat{S}_F^n$, the capacity region remains unchanged.

Claim 2. Any achievable $(R_F, R_N)$ is included in the capacity region of a BEC with $\hat{S}_F^n$ instead of $S_F^n$ while other parameters are kept the same and vice versa.

Proof. Let $I(W_F; Y_F^n | S_F^n) = I(W_F; Y_F^n | S_F^n, K^n)$

- $\min \{\delta, 1 - \delta\} = \delta$ : In this case, to maximize $\Pr \left( \hat{S}_F[t] = S_N[t] = 0 \right)$, the transition from $S_F^n$ to $\hat{S}_F^n$ would be such that to minimize the times in which $\hat{S}_F[t] = 1$ and $S_N[t] = 0$ by setting $D_{10} (1 - \delta) n$ ones to zeros when $S_F[t] = 1$ and $S_N[t] = 0$. The transition from $\hat{S}_F^n$ to $\hat{S}_F^n$ would be such that to minimize the times in which $\hat{S}_F[t] = 1$ and $S_N[t] = 1$ by changing $D_{10} (1 - \delta) n$ zeros in $\hat{S}_F^n$ to ones when $S_F[t] = 0$ and $S_N[t] = 1$. In effect, such transition from $S_F^n$ to $\hat{S}_F^n$, increases the probability of $\hat{S}_F[t] = S_N[t] = 0$ by $D_{10} (1 - \delta)$, and since $D_{10} \leq D_{FB}$, the maximum occurs at $D_{10} = D_{FB}$. Based on Figure 3, the discussion above can be explained as follows. The transition from $S_F^n$ to $\hat{S}_F^n$ removes realizations from state (b) and adds them to state (d), while transition from $\hat{S}_F^n$ to $\hat{S}_F^n$ moves realizations from state (c) to (d). Thus, end-to-end transitions are from state (b) to (c).

\[
\min \{\delta, 1 - \delta\} = \delta
\]

- $\min \{\delta, 1 - \delta\} = 1 - \delta$ : In this case, to maximize $\Pr \left( \hat{S}_F[t] = S_N[t] = 0 \right)$, the transition from $S_F^n$ to $\hat{S}_F^n$ would be such that to minimize the times in which $\hat{S}_F[t] = 1$ and $S_N[t] = 0$ by changing $D_{01} \delta n$ zeros to ones when $S_F[t] = 1$ and $S_N[t] = 1$. The transition from $\hat{S}_F^n$ to $\hat{S}_F^n$ would be such that to minimize the times in which $\hat{S}_F[t] = 1$ and $S_N[t] = 0$ by changing $D_{01} \delta n$ ones in $\hat{S}_F^n$ to zeros when $S_F[t] = 1$ and $S_N[t] = 0$. In effect, such transition from $S_F^n$ to $\hat{S}_F^n$, increase the probability of $\hat{S}_F[t] = S_N[t] = 0$ by $D_{01} \delta$, and since $D_{01} \leq D_{FB}$, the maximum occurs at $D_{01} = D_{FB}$.

To prepare for the final step, we define

\[
\zeta_{out} \triangleq \frac{D_{FB}}{\min \{\delta, 1 - \delta\}}.
\]
Based on this definition and [14], we have
\[
\min_{D_{10} + D_{01} \leq D_{FB}} \Pr \left( \{ \tilde{S}_F[t] = S_N[t] = 0 \}^c \right) = \gamma_{out} (1 - \delta) + (1 - \gamma_{out}) (1 - \delta^2) .
\] (16)

Further, from rate-distortion theory, we have \( D_{FB} \geq D^* \). Thus, for any feedback strategy, we have
\[
A \triangleq \max_{D_{FB} \geq D^*} \min_{D_{10} + D_{01} \leq D_{FB}} \Pr \left( \{ \tilde{S}_F[t] = S_N[t] = 0 \}^c \right) = \gamma_{out} (1 - \delta) + (1 - \gamma_{out}) (1 - \delta^2) ,
\] (17)
where \( \gamma_{out} \) is defined in [9], and this bound is attained for a code that achieves the rate-distortion bound. In other words, for any feedback strategy, we have
\[
\Pr \left( \{ \tilde{S}_F[t] = S_N[t] = 0 \}^c \right) \leq \gamma_{out} (1 - \delta) + (1 - \gamma_{out}) (1 - \delta^2)
\] (18)
where \( \tilde{S}_F^n \) satisfies the conditions in [11].

**Step 4:** For the final step, we first prove the following result.

**Claim 4.** For the two-user BEC with rate-limited feedback as described in Section[11] and for any input distribution, we have
\[
H (Y^n_F | W_N, S^n_F, S^n_N) - \beta_{out} H (Y^n_N | W_N, S^n_F, S^n_N) \leq 0,
\] (19)
where \( \beta_{out} \) is given in [8].

**Proof.**
\[
H (Y^n_N | W_N, S^n_F, S^n_N) = H (Y^n_N | W_N, S^n_F) + I (Y^n_N; S^n_F | W_N)
\]
\[
\leq H (Y^n_N | W_N, S^n_F, K^n, S^n_N) + I (Y^n_N; S^n_F | W_N, K^n, S^n_N)
\]
\[
= H (Y^n_N | W_N, S^n_F, K^n, S^n_N) + I (\tilde{S}_F^n | W_N, K^n, S^n_N)
\]
\[
\geq H (Y^n_N | W_N, \tilde{S}_F^n, K^n, S^n_N) + I (\tilde{S}_F^n | Y^n_N, W_N, S^n_N)
\]
\[
\geq \sum_{t=1}^n (1 - \delta) H (Y^n_N | W_N, \tilde{S}_F^t, S^t_N)
\]
\[
\geq \sum_{t=1}^n (1 - \delta) H (Y^n_N | W_N, \tilde{S}_F^t, S^t_N)
\]
\[
\geq \sum_{t=1}^n (1 - \delta) H (\tilde{Y}_F^t | Y^n_N, W_N, \tilde{S}_F^t, S^t_N)
\]
\[
\geq \sum_{t=1}^n \frac{1}{A} H (\tilde{Y}_F^t | Y^n_N, W_N, \tilde{S}_F^t, S^t_N)
\]
where (a) & (b) hold since conditioned on \( K^{t-1} \), \( X[t] \) is independent of all other channel parameters; (c) follows the fact that \( \tilde{S}_F^n \) results in \( K^n \); (d) follows from the chain rule and the causality of the channel; (e) holds since \( S_N[t] \) is a Bernoulli \( (1 - \delta) \) process; (f) is true since \( X[t] \) is independent of channel realizations at time instant \( t \); (g) holds since conditioning reduces entropy; (h) follows by the definition of \( A \) in [17], normalizing by the probability that at least one signal is not erased, and ensuring the inequality holds for any feedback strategy; (i) follows the causality assumption; and (j) holds as the discrete entropy function is non-negative.

Finally, we are ready to prove the outer-bounds.
\[
n (R_F + \beta_{out} R_N) = H (W_F) + \beta_{out} H (W_N)
\]
\[
= H (W_F | W_N, S^n_F, S^n_N) + \beta_{out} H (W_N | S^n_F, S^n_N)
\]
\[
\leq I (W_F; Y^n_N | W_N, S^n_F, S^n_N) + \beta_{out} I (W_N; Y^n_N | S^n_F, S^n_N) + n \epsilon_n
\]
\[
= H (Y^n_N | W_N, S^n_F, S^n_N) - H (Y^n_N | W_F, W_N, S^n_F, S^n_N)
\]
\[
\geq \beta_{out} H (Y^n_N | S^n_F, S^n_N) - \beta_{out} H (Y^n_N | W_N, S^n_F, S^n_N) + n \epsilon_n
\]
Claim \[8\] \[
\leq \beta_{out} H (Y^n_N | S^n_F, S^n_N) + n \epsilon_n \leq n \beta_{out} (1 - \delta) + n \epsilon_n ,
\]
where \( \epsilon_n \to 0 \) as \( n \to \infty \). Dividing both sides by \( n \) and taking the limit for \( n \to \infty \) completes the proof. The other outer-bound in Theorem[1] can be obtained similarly.

**ACKNOWLEDGEMENT**

The research of A. Vahid was supported in part by NSF grant ECCS-2030285.
REFERENCES

[1] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y.-H. Kim, “Wire-tap channel with secure rate-limited feedback,” *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, 2009.

[2] O. Shayevitz and M. Wigger, “On the capacity of the discrete memoryless broadcast channel with feedback,” *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1329–1345, 2012.

[3] S. B. Amor, Y. Steinberg, and M. Wigger, “MIMO MAC-BC duality with linear-feedback coding schemes,” *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 5976–5998, 2015.

[4] Y. Wu and M. Wigger, “Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels,” *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2009–2033, 2016.

[5] A. Vahid, I.-H. Wang, and S.-C. Lin, “Capacity results for erasure broadcast channels with intermittent feedback,” in 2019 *IEEE Information Theory Workshop (ITW)*, pp. 1–5, IEEE, 2019.

[6] A. Vahid, S.-C. Lin, I. Wang, et al., “Erasure broadcast channels with intermittent feedback,” *arXiv preprint arXiv:2005.14383*, 2020.

[7] C. He and S. Yang, “On the two-user erasure broadcast channel with one-sided feedback,” in 2017 *IEEE Information Theory Workshop (ITW)*, pp. 494–498, IEEE, 2017.

[8] S.-C. Lin and I.-H. Wang, “Gaussian broadcast channels with intermittent connectivity and hybrid state information at the transmitter,” *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6362–6383, 2018.

[9] S.-C. Lin, I.-H. Wang, and A. Vahid, “No feedback, no problem: Capacity of erasure broadcast channels with single-user delayed CSI,” in 2019 *IEEE International Symposium on Information Theory (ISIT)*, pp. 1647–1651, IEEE, 2019.

[10] S.-C. Lin, I.-H. Wang, and A. Vahid, “Capacity of broadcast packet erasure channels with single-user delayed CSI,” submitted to *IEEE Transactions on Information Theory*, 2020.

[11] A. Vahid and A. S. Avestimehr, “The two-user deterministic interference channel with rate-limited feedback,” *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, July 2010.

[12] A. Vahid, C. Suh, and A. S. Avestimehr, “Interference channels with rate-limited feedback,” *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2788–2812, 2012.

[13] M. Gastpar, A. Lapidoth, Y. Steinberg, and M. Wigger, “Coding schemes and asymptotic capacity for the Gaussian broadcast and interference channels with feedback,” *IEEE transactions on information theory*, vol. 60, no. 1, pp. 54–71, 2013.

[14] C. Karakus, I.-H. Wang, and S. Diggavi, “Gaussian interference channel with intermittent feedback,” *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4663–4699, 2015.

[15] D. Tuninetti, “An outer bound for the memoryless two-user interference channel with general cooperation,” in 2012 *IEEE Information Theory Workshop*, pp. 217–221, IEEE, 2012.

[16] Z. Cheng and N. Devroye, “Two-way networks: when adaptation is useless,” *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1793–1813, 2013.

[17] C. Suh, J. Cho, and D. Tse, “Two-way interference channel capacity: How to have the cake and eat it too,” *IEEE Transactions on Information Theory*, vol. 64, no. 6, pp. 4259–4281, 2018.

[18] A. Vahid, V. Aggarwal, A. S. Avestimehr, and A. Sabharwal, “On the capacity of multi-hop wireless networks with partial network knowledge,” in 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 1030–1037, IEEE, 2010.

[19] V. Aggarwal, A. S. Avestimehr, and A. Sabharwal, “On achieving local view capacity via maximal independent graph scheduling,” *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2711–2729, 2011.

[20] A. Vahid and R. Calderbank, “Impact of local delayed CSIT on the capacity region of the two-user interference channel,” in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2421–2425, IEEE, 2015.

[21] A. Vahid and R. Calderbank, “Two-user erasure interference channels with local delayed CSIT,” *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 4910–4923, 2016.

[22] A. Vahid and R. Calderbank, “When does spatial correlation add value to delayed channel state information?,” in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2624–2628, IEEE, 2016.

[23] A. Vahid, V. Aggarwal, A. S. Avestimehr, and A. Sabharwal, “Interference management with mismatched partial channel state information,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, pp. 1–22, 2017.

[24] A. Vahid and R. Calderbank, “Throughput region of spatially correlated interference packet networks,” *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 1220–1235, 2019.

[25] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, “Binary fading interference channel with delayed feedback,” in 2012 *IEEE International Symposium on Information Theory Proceedings*, pp. 1882–1886, IEEE, 2012.

[26] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, “Capacity results for binary fading interference channels with delayed CSIT,” *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6093–6130, 2014.

[27] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, “Communication through collisions: Opportunistic utilization of past receptions,” in *INFOCOM*, pp. 2553–2561, IEEE, 2014.

[28] A. Vahid and R. Calderbank, “ARQ for interference packet networks,” in 2018 *IEEE International Symposium on Information Theory (ISIT)*, pp. 781–785, IEEE, 2018.

[29] M. Gatzianas, L. Georgiadis, and L. Tassiulas, “Multituser broadcast erasure channel with feedback–capacity and algorithms,” *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5779–5804, 2013.