Capacity and Coding Schemes for Binary Stochastic-Adversarial Channels with Feedback Snooping

Vinayak Suresh, Eric Ruzomberka, Chih-Chun Wang and David J. Love

Abstract

With the advent of 5G and technologies such as cloud computing, Internet-of-Things (IoT), etc, future communication networks will consist of a large number of heterogeneous devices connected together. A critical aspect will be ensuring that communication is not only fast and reliable, but also secure. Motivated by this, we consider the problem of communicating a message reliably across a binary erasure channel (BEC(q)) or a binary symmetric channel (BSC(q)) against an adversary actively injecting additional erasures or flips at the channel's input. The adversary has a total error budget equal to a fixed fraction p of the codeword length and knows the transmission scheme agreed upon by the communicating terminals. Further, he has the capability to causally snoop in on both the transmitter and the receiver in real time, i.e., if $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ denote the transmitted and the received codewords respectively, at each time k, he knows $(x_1, x_2, \cdots, x_k)$ and $(y_1, y_2, \cdots, y_{k-1})$. The adversary is free to employ any attack using his side-information that respects his budget constraint. We prove an information-theoretic tight capacity characterization as a function of p and q for (i) the erasure adversary with a BEC(q) and (ii) the bit-flip adversary with a BSC(q). A unique feature of our models is the compounding of stochastic and adversarial noise sources. Our analysis reveals the worst-case adversarial attacks for both models and proves the existence of coding schemes that achieve rates equal to the capacity for any adversarial attack. In the case of bit-flips, we show that, interestingly, when p is below a certain threshold (that depends on q), the adversary is no worse than an i.i.d. memory-less noise source.

This work was funded in part by the National Science Foundation under grants CNS1642982, CCF1816013, CCF2008527 and EEC1941529. A preliminary version was presented at the 2021 IEEE International Symposium on Information Theory (ISIT) [1] with an extended version of the conference paper available at [2].

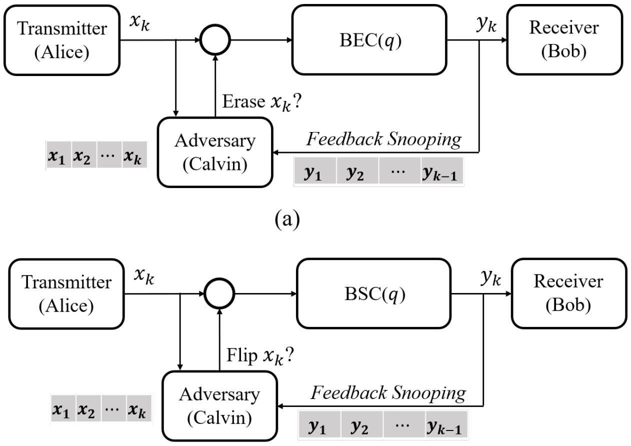
V. Suresh, E. Ruzomberka, C.-C. Wang and D. J. Love are with the Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, USA. (email: suresh20@purdue.edu, eruzombe@purdue.edu, chihw@purdue.edu, djlove@purdue.edu)

I. Introduction

Due to a massive increase in the number of devices connected together, security is a re-emerging concern for wireless networks. There is a push from 3GPP, government and other stakeholders to adopt zero-trust design principles for 5G networks and beyond [3], securing systems from attackers both outside and within the network. From the zero-trust perspective, designers must assume an *open network* where all network links can be intercepted by an attacker. Furthermore, due to the ease-of-access of the wireless medium, designers must plan for attacks at the physical layer, including denial-of-service (DoS) attacks (i.e., jamming) or other attacks which can lead to network wide security vulnerabilities. Such attacks can come from untrusted devices or compromised hardware (i.e., hardware Trojans) [4]–[6]. Trojans can use real time information snooped from a link to design optimal attacks on error control systems – systems which are not currently designed to defend against these attacks. In this article, we develop secure error control coding techniques against such threats.

Specifically, consider the following situation depicted in Fig. 1. Alice wishes to communicate a message reliably to Bob over a binary erasure channel (BEC(q)) or a binary symmetric channel (BSC(q)) in the presence of Calvin, who can introduce additional noise at the channel's input by erasing or flipping bits. Calvin assumes the role of an *online* adversary who has the ability to spy on *both* terminals in *real time*. He may only impact a certain number of bits but can otherwise freely corrupt parts of the transmission. Here, his budget is specified as a fraction of the codeword length (pn erasures or flips where n is the codeword length). What is the largest rate at which reliable communication is possible (i.e. *channel capacity*) in this setting? Answering this question is the central goal of this paper.

Many of the channel models in information theory are broadly of two kinds. On one side are *stochastic* models whose behavior is characterized by a probability law and errors get injected independent of the communication scheme. Here, it is sufficient to deal with *average-case* errors. On the other extreme are *adversarial* models where one must deal with the *worst-case* errors. As expected, the latter often behave much differently from the former. In the case of adversarial channels, the capacity generally depends strongly on what the adversary knows. An *oblivious* adversary [7]–[11] is one who possibly knows the coding scheme agreed upon by Alice and Bob but has no knowledge of the transmitted codeword. In complete contrast is the *omniscient* adversary [12]–[14] who non-causally knows the entire length-n codeword chosen by Alice for



(a)

(b)

Fig. 1. Channel models considered in this work - (a) BEC(q)-ADV(p)-FS and (b) BSC(q)-ADV(p)-FS. Calvin, who at each time k knows $(x_1, x_2, \cdots, x_k)$ and $(y_1, y_2, \cdots, y_{k-1})$, and also knows the transmission scheme being employed by Alice and Bob, is constrained such that he may only inject up to pn erasures or flips in total.

transmission. An intermediate model also considered in this paper is that of an *online* or *causal* adversary [15]–[18] wherein at any point during the transmission, the adversary has access to part of the codeword that is transmitted thus far, i.e., if $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is the codeword transmitted, Calvin at each time k knows $(x_1, x_2, \cdots, x_k)$. Another interesting set of models are the *delayed* adversary model [19], [20] and the *look-ahead* adversary model [15] where Calvin at each time k knows $(x_1, x_2, \cdots, x_{k+d \cdot n})$, where d is the delay ($d < 0$) or the look-ahead ($d > 0$) parameter. Different from these is also the *myopic* adversary model [21], [22] where Calvin knows only a noisy version of $\mathbf{x}$.

Along with the adversary's side information, another important criterion that affects the capacity is whether Alice and Bob have any shared randomness between them that is unknown to Calvin. In most cases, it turns out that the adversary in these settings is no worse than an i.i.d. memory-less noise source [15], [23]–[25]. Moreover, from a practical perspective, today's wireless systems do not have physical layers based on such type of shared randomness. Therefore in this paper, we do not allow any shared randomness between the terminals. However, we allow Alice to employ *stochastic encoding* or randomized encoding using private random coins that are shared neither to Bob nor Calvin.

Without Calvin's presence, i.e. when $p = 0$, our models reduce to the classical BEC(q) or the BSC(q). When there is no random channel present, i.e., $q = 0$, the only source of noise is adversarial for which a complete capacity characterization is given by [15]–[17]. Our models differ from the ones considered previously in two ways:

- **Mixing of random and adversarial noise**: From Fig. 1, the noise in the received word is affected by the random channel BEC/BSC as well as the actions of Calvin who is erasing/flipping bits. For example in the erasure case, a bit not erased by Calvin can be erased by the BEC. In the bit-flip case, the situation is harder because a bit flipped by Calvin may even get "unflipped" by the BSC. Conceptually, we think of the stochastic channel as the main channel through which Alice and Bob communicate, and Calvin as a malicious entity who attempts to actively disrupt the transmission. Since we only deal with binary channels, we refer to our models as *binary stochastic-adversarial channels*. Study of real input-output channels such as the AWGN channel are left for future investigation.
- **Feedback to adversary**: In our setting, we will allow Calvin access to Bob's reception through *feedback snooping*, as shown in Fig. 1. This becomes important due to the presence of the stochastic channel that also influences the bits received at Bob. Note that feedback snooping is unnecessary when $q = 0$.

We note that our models are in fact special cases of the more general framework of arbitrarily varying channels (AVCs) [7], [26]. However, known results for AVCs do not imply the results of this paper and therefore we do not pursue this connection. Our contributions are briefly summarized as follows:

- We provide a complete characterization of capacity in the case of erasures ($C^{erase}(p, q)$) for arbitrary budget parameter $p \in [0, 1]$ and erasure probability $q \in [0, 1]$. Our result implies that the presence of the random channel BEC(q) in addition to causal adversarial erasures scales the capacity expression of the $q = 0$ case by a multiplicative factor.
- We also provide a complete capacity characterization in the case of bit-flips ($C^{flip}(p, q)$) for arbitrary budget parameter $p \in [0, 1]$ and flip probability $q \in [0, 1/2]$. We show that for every $q \in [0, 1/2)$, there is a threshold $p_q > 0$ s.t. when $p < p_q$, Calvin can do no better than making flip decisions in an i.i.d. manner. In other words, an adversary when weak enough is no worse than an i.i.d. memory-less noise source. Here, $p_q \to 0$ as $q \to 1/2$.
- For each model, we characterize the worst-case adversarial attacks and prove the existence of coding schemes that allow Alice to transmit reliably at rates arbitrarily close to capacity, no matter the adversary's strategy.

A preliminary version of this work was presented at the 2021 IEEE International Symposium on Information Theory [1]. An extended version of the ISIT conference paper with longer proofs is available at [2]. In [1], [2], while the capacity for the erasure model was completely characterized, only upper and lower bounds were given for the harder bit-flip model. In this work, we close this gap and show that the converse sketched in [1], [2] is in fact tight. Inclusion of secrecy constraints where Alice must not only convey her message reliably to Bob but also hide it from Calvin, is not considered here and left for future investigation. The rest of the paper is organized as follows. Section II formally defines the channel models and the capacity characterization problem. In Section III, we state our main capacity results. Converse proofs are provided in Section IV and proofs for achievability are provided in Section V. Finally, conclusions and possible future research directions are discussed in Section VI.

II. Preliminaries

A. Channel Models

The channel models are depicted in Fig.1. Encoding is done over a fixed block-length of n channel uses, and the size of the message set at the transmitter is 2^{nR}. Consider first the model for the case of erasures. Alice (the transmitter) attempts to convey a message to Bob (the receiver) over a BEC(q), in the presence of a p-limited causal adversary (Calvin) where the terms will be clarified shortly. The input and output alphabets are $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, 1, \Lambda\}$ respectively, where Λ denotes an erasure symbol. We allow stochastic encoding and assume the presence of local randomness available *only* to Alice for this purpose. Denote $x_k \in \mathcal{X}$ to be the bit selected by the transmitter at channel use k. At time k, Calvin makes a decision on whether to erase x_k based on his side-information to be specified. If Calvin erases x_k, the received symbol at time k at the receiver is an erasure, i.e., $y_k = \Lambda$. If Calvin decides not to erase x_k, then x_k is erased with probability q, i.e., $y_k = x_k$ with probability $1 - q$ and $y_k = \Lambda$ with probability q. We now specify the side-information available to Calvin:

- **Knowledge of transmission scheme:** Calvin has knowledge of the transmission scheme agreed upon by Alice and Bob.
- **Transmitter snooping:** Calvin has causal access to symbols being transmitted by Alice, i.e., at each channel use k, $1 \leq k \leq n$, Calvin knows $(x_1, x_2, \cdots, x_k) \in \mathcal{X}^k$.

- **Feedback snooping:** Calvin has the capability to spy into Bob's reception through a noise-free *strictly* causal feedback link as shown in Fig. 1. At each channel use k, $1 \leq k \leq n$, Calvin knows $(y_1, y_2, \cdots, y_{k-1}) \in \mathcal{Y}^{k-1}$.

Thus, Calvin's decision on whether or not to erase x_k is a function of the transmission rule, $(x_1, x_2, \cdots, x_k) \in \mathcal{X}^k$, and $(y_1, y_2, \cdots, y_{k-1}) \in \mathcal{Y}^{k-1}$. A power constraint is further imposed by enforcing Calvin to be p-limited, meaning that he can erase at most a constant fraction p of the bits, i.e., if $\mathbf{a} \in \{0,1\}^n$ denotes the positions where Calvin decides to erase symbols from $(x_1, x_2, \cdots, x_n)$, we must have $weight(\mathbf{a}) \leq pn$. We refer to this model as *the BEC causal adversarial channel with feedback snooping* (or BEC(q)-ADV(p)-FS). Note that the BEC block in Fig. 1(a) is different from the classical BEC. If Calvin erases x_k to an erasure symbol Λ, we have $y_k = \Lambda$, where Λ does not carry any information.

We also consider a related and more interesting model (Fig. 1(b)) where Calvin can attempt to flip up to pn bits and the stochastic channel is a BSC(q) instead of a BEC(q). The input and output alphabets are revised to $\mathcal{X} = \{0,1\}$ and $\mathcal{Y} = \{0,1\}$. At time k, Calvin produces $a_k \in \mathcal{A} = \{0,1\}$ based on his side information which is the same as that for erasures, i.e., at time k, he knows $(x_1, x_2, \cdots, x_k)$, the transmission scheme, and $(y_1, y_2, \cdots, y_{k-1})$. The received symbol at time k at the receiver is

$$y_k = \begin{cases} x_k \oplus a_k \oplus 1 & \text{with prob. } q \\ x_k \oplus a_k & \text{with prob. } 1-q \end{cases},$$

where $\oplus$ denotes mod-2 addition and $q \in [0, 1/2]$. Hence, $\mathbf{a} \in \{0,1\}^n$ denotes the positions where Calvin injects bit-flips and the constraint on the adversary can be expressed as $weight(\mathbf{a}) \leq pn$. Note that a flip-attempt of Calvin can now be undone by the BSC. This happens exactly at positions where both Calvin and the BSC inject errors. This is in contrast to the case of erasures where a bit erased by Calvin remains erased. This model is referred to as *the BSC causal adversarial channel with feedback snooping* (or BSC(q)-ADV(p)-FS).

Our aim is to characterize the capacity of these channels, i.e., the largest value of R such that Alice can reliably convey one out of 2^{nR} possible messages to Bob. The capacities of the BEC(q)-ADV(p)-FS channel and the BSC(q)-ADV(p)-FS channel are denoted by $C^{erase}(p,q)$ and $C^{flip}(p,q)$ respectively. Precise definitions to follow.

Definitions: The transmitted message is denoted by the random variable (r.v.) $\mathbf{U}$ chosen uniformly from the message set $\mathcal{U} = \{1, 2, 3, \cdots, 2^{nR}\}$. The Hamming distance between $\mathbf{w}$

and $\mathbf{z}$ will be denoted by $d_H(\mathbf{w},\mathbf{z})$. We denote by $\mathcal{C}(n,R)$ a code of rate R and block-length n. A deterministic code $\mathcal{C}=(\Phi_d,\Gamma_d)$ consists of an encoder map $\Phi_d:\mathcal{U}\rightarrow\mathcal{X}^n$ and a decoder map $\Gamma_d:\mathcal{Y}^n\rightarrow\mathcal{U}$, where each message is associated to a unique codeword. In case of stochastic encoding, a codeword $\mathbf{x}$ is selected for a message u according to a chosen conditional distribution $\tilde{\Phi}(\cdot|u)$ defined on $\mathcal{X}^n$. A stochastic code $\mathcal{C}=(\tilde{\Phi},\Gamma)$ is fully specified by defining all conditional distributions $\left\{\tilde{\Phi}(\cdot|u)\right\}_{u\in\mathcal{U}}$ and decoder $\Gamma:\mathcal{Y}^n\rightarrow\mathcal{U}$. Without loss of generality, we assume in proving converse results that no two distinct messages map to the same codeword.

Denote the transmitted and received codewords by $\mathbf{x}$ and $\mathbf{y}$ respectively. A strategy $\mathfrak{S}$ for Calvin consists of (possibly stochastic) maps $g_1,g_2,\cdots,g_n$, where his error injections are given by

$$a_k=g_k(\mathcal{C},\mathbf{x}_1^k,\mathbf{y}_1^{k-1})\quad k=1,2,\cdots,n.$$

Equivalently, for each k, a_k is a Bernoulli random variable whose success probability is a function only of the transmission rule $\mathcal{C}$, $\mathbf{x}_1^k$ and $\mathbf{y}_1^{k-1}$. $\mathfrak{S}$ is feasible only if for every $\mathcal{C}$, $\mathbf{x}$ and $\mathbf{y}$, $weight((a_1,a_2,\cdots,a_n))\leq pn$ holds almost surely. The set of all feasible strategies for Calvin is denoted by ADV(p). The (maximum) probability of error is then defined as

$$P_e(\tilde{\Phi},\Gamma)=\max_{u\in\mathcal{U}}\max_{\mathfrak{S}\in\text{ADV}(p)}\sum_{\mathbf{y}}\sum_{\mathbf{x}}P(\mathbf{y}|\mathbf{x},\mathfrak{S})\tilde{\Phi}(\mathbf{x}|u)\mathcal{I}(\Gamma(\mathbf{y})\neq u)\tag{1}$$

where $\mathcal{I}(.)$ denotes the indicator function.

When proving achievability results, we consider for analytical simplicity the following alternate view of a stochastic code. Alice is endowed with a set $\mathcal{S}$ of private secrets or keys and the stochastic code is defined by a deterministic map $\Phi:\mathcal{U}\times\mathcal{S}\rightarrow\mathcal{X}^n$. For a given message $u\in\mathcal{U}$, the codeword $\Phi(u,s)$ is selected by picking a secret $s\in\mathcal{S}$ uniformly randomly. As discussed in [16], this definition is essentially equivalent and does not change the capacity. In this case, the (maximum) probability of error from (1) is revised to $P_e(\Phi,\Gamma)=\max_{u\in\mathcal{U}}\max_{\mathfrak{S}\text{ADV}(p)}\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\sum_{\mathbf{y}}P(\mathbf{y}|\Phi(u,s),\mathfrak{S})\mathcal{I}(\Gamma(\mathbf{y})\neq u)$. The probability of decoding error is averaged over all possible secrets available to Alice for encoding.

Rate $R>0$ is achievable if for every $\epsilon>0$, there is a sequence of rate $R-\epsilon$ codes of increasing block-lengths $\{\mathcal{C}(n,R-\epsilon)\}_{n\geq1}$ such that for any $\delta>0$, there is an N so that $P_e(\mathcal{C}(n,R-\epsilon))<\delta$ for any $n>N$. Capacity is defined to be the supremum of all achievable rates. For $x,y\in[0,1/2]$, define $x\star y=x(1-y)+y(1-x)$. Note that $x\star y=1/2$ iff either $x=1/2$ or $y=1/2$ (or both).

B. Simple Converse Bounds - The I.I.D. Attack

We begin with simple converse bounds for both channel models. These follow from the following adversarial attack for Calvin - he ignores his side information completely and simulates an i.i.d. memory-less noise source while respecting his budget constraint.

Lemma 1. *The capacity $C^{erase}(p,q)$ of the BEC(q)-ADV(p)-FS channel satisfies the bound*

$$C^{erase}(p,q) \leq (1-p)(1-q). \tag{2}$$

Proof. For $\delta > 0$, Calvin erases each bit x_i independently with probability $p-\delta$. By the Chernoff bound, he does not exceed his budget with probability at least $1-2^{-\Omega(\delta^2 n)}$. Since the combination of this attack with the $BEC(q)$ is the $BEC(s)$ with $s = p-\delta+q-(p-\delta)q$, the capacity is bounded as $C^{erase}(p,q) \leq (1-p+\delta)(1-q)$. Letting $\delta \to 0$ completes the proof. □

Lemma 2. *The capacity $C^{flip}(p,q)$ of the BSC(q)-ADV(p)-FS channel satisfies the bound*

$$C^{flip}(p,q) \leq 1-h_2(p\star q). \tag{3}$$

Proof. For $\delta > 0$, Calvin flips each bit x_i independently with probability $p-\delta$, staying within his budget with probability at least $1-2^{-\Omega(\delta^2 n)}$. Since the effective channel is $BSC((p-\delta)\star q)$, the capacity is bounded as $C^{flip}(p,q) \leq 1-h_2\left((p-\delta)\star q\right)$. Finally, let $\delta \to 0$ to get (3). □

We will show that for the BEC(q)-ADV(p)-FS channel, the i.i.d. erasure attack in Lemma 1 is *always* sub-optimal, as one would expect. In contrast however, for the BSC(q)-ADV(p)-FS channel, there are regimes where the i.i.d. bit-flip attack in Lemma 2 is optimal. Here, the side information available to Calvin as specified in II-A proves to be of no benefit, and Calvin is no worse than an i.i.d. Bernoulli memory-less noise source.

C. Effective number of erasures or flips

By the Chernoff bound, the BEC(q)/BSC(q) when acting alone (i.e., with no adversary) induces about qn erasures/flips. In our set-up, we also have Calvin who can introduce up to pn additional erasures/flips. However, since Calvin is causal, his error pattern and the error pattern induced by the random channel may have several overlapping error injections. The total number of errors will thus be much less than $pn+qn$. Consider the following lemma.

Lemma 3. *Let $X_1, X_2, \cdots, X_n$ be i.i.d. Ber(q) indicator random variables representing the erasure sequence injected by a BEC(q). Let $Y_1, Y_2, \cdots, Y_n$ be indicator random variables where, for each j, Y_j is Bernoulli distributed with a success probability that is possibly a function of $X_1, X_2, \cdots, X_{j-1}, Y_1, Y_2, \cdots, Y_{j-1}$, subject to the constraint that the random variable $\sum_j Y_j$ is almost surely less than or equal to pn. For $\delta > 0$, defining the event*

$$E = \left\{ \sum_{j=1}^{n} \mathcal{I}(X_j = 1 \textit{ or } Y_j = 1) \leq (p + q - pq)n + \delta n \right\},$$

we have $P(E) \geq 1 - 2^{-\Omega(\delta^2 n)}$.

Proof. By defining a suitable martingale, the proof is a simple application of Azuma's inequality (e.g. [27]). Let $Z_j = \mathcal{I}(X_j = 1 \textit{ or } Y_j = 1)$. Define for $j = 1, 2, \cdots, n$, $P_j = Z_j - \mathbb{E}(Z_j \mid X_1, X_2, \cdots, X_{j-1}, Y_1, Y_2, \cdots, Y_{j-1})$, and $S_j = \sum_{k=1}^{j} P_k$. Clearly, S_j is a martingale because $\mathbb{E}(S_{j+1} \mid X_1, \cdots, X_j, Y_1, \cdots, Y_j) = S_j$. Note that $|S_j - S_{j-1}| = |P_j| \leq 1$ holds almost surely. Thus, by Azuma's inequality, $Pr\left(|S_n| \geq \delta n\right) \leq 2e^{-\frac{\delta^2 n}{2}}$. The required result then follows from the constraint $\sum_j Y_j \leq pn$ and the fact that for each j, X_j and Y_j are independent. □

Using arguments similar to the proof of Lemma 3, we can also show the following Lemma.

Lemma 4. *Let $X_1, X_2, \cdots, X_n$ be i.i.d. Ber(q) random variables representing the error sequence injected by a BSC(q). Let $Y_1, Y_2, \cdots, Y_n$ be random variables where, for each j, Y_j is Bernoulli distributed with a success probability that is possibly a function of $X_1, \cdots, X_{j-1}, Y_1, \cdots, Y_{j-1}$, subject to the constraint $\sum_j Y_j \leq pn$ almost surely. For $\delta > 0$, defining the event $E = \left\{ \sum_{j=1}^{n} (X_j \oplus Y_j) \leq (p \star q)n + \delta n \right\}$, we have $P(E) \geq 1 - 2^{-\Omega(\delta^2 n)}$.*

Let $\delta > 0$ be a small arbitrary constant. Lemmas 3 and 4 imply that under any strategy employed by Calvin, we have the following:

- For the BEC(q)-ADV(p)-FS channel, the total effective number of erasures injected on to the received codeword due to actions of both Calvin and the BEC(q) does not exceed $(p + q - pq + \delta)n$ with probability at least $1 - 2^{-\Omega(\delta^2 n)}$.
- For the BSC(q)-ADV(p)-FS channel, the total effective number of flips injected on to the received codeword due to actions of both Calvin and the BSC(q) does not exceed $(p \star q + \delta)n$ with probability at least $1 - 2^{-\Omega(\delta^2 n)}$.

The above results imply that insofar as the total effective number of flips or erasures is concerned, Calvin cannot use his side information to improve over an i.i.d. attack. Lemmas 3 and 4 will also be essential to proving our achievability results.

III. Main Results

A. Results for Erasures

Theorem 1. *The capacity $C^{erase}(p,q)$ of the BEC(q)-ADV(p)-FS channel is*

$$C^{erase}(p,q)=\begin{cases}(1-2p)(1-q), & 0\le p\le \frac{1}{2},\ 0\le q\le 1\\ 0, & \textit{otherwise}\end{cases}. \tag{4}$$

When there is no BEC, i.e., when $q=0$, our model reduces to the one studied in [15], [16]. Our result implies that in the setting where both causal adversarial erasures and random erasures are present, the capacity expression is scaled by a factor of $(1-q)$.

B. Results for Bit-flips

Theorem 2. *For $p\in[0,1/4]$ and $q\in[0,1/2]$, the capacity $C^{flip}(p,q)$ of the BSC(q)-ADV(p)-FS channel is*

$$C^{flip}(p,q)=\min_{x\in[0,p]}\alpha(p,x)\left(1-h_2\left(\frac{x(1-2q)}{\alpha(p,x)}+q\right)\right) \tag{5}$$

where $\alpha(p,x)=1-4(p-x)$. If $p\ge 1/4$, we have $C^{flip}(p,q)=0$.

When $q=0$, i.e., there is no BSC, the channel model reduces to that considered in [15], [16], and the capacity expression (5) matches with the result proved in [15], [16]. As shown in Appendix B, the solution $C^{flip}(p,q)$ to the optimization problem in (5) for any fixed $q\in[0,1/2)$ is

$$C^{flip}(p,q)=\begin{cases}1-h_2(p\star q) & 0\le p\le p_q\\ \frac{1-4p}{1-4p_q}\left(1-h_2\left(p_q\star q\right)\right) & p_q\le p\le 1/4\\ 0 & p\ge 1/4\end{cases},$$

where p_q is the unique solution in $(0,1/2)$ of the equation

$$4+(1+2q)\log_2\left(p_q\star q\right)+(3-2q)\log_2\left(1-p_q\star q\right)=0. \tag{6}$$

In Fig. 2, $C^{flip}(p,q)$ is plotted as a function of p for various values of q, specifically, $q=0.0,0.1,0.2$. We make the following observations:

- From [15], [16], $C^{flip}(p, 0) > 0$ for $p \in [0, 1/4)$. Here, we have $C^{flip}(p, q) > 0$ for all $q \in [0, 1/2)$ and $p \in [0, 1/4)$. Thus, the addition of the BSC stochastic channel does not change the support over p for which the a positive rate is achievable.
- For $0 \leq p \leq p_q$, $C^{flip}(p, q)$ is convex and equal to $1 - h_2(p \star q)$. This implies that when $0 \leq p \leq p_q$, the i.i.d. bit-flip attack strategy in Section II-B is optimal for the adversary. In this regime, the knowledge of the encoding scheme or the ability to spy on Alice or Bob buys Calvin no benefit.
- Solving (6), it can be seen, as $q \searrow 0$, $p_q \nearrow p_0 = \frac{1}{6}\left(5 - \frac{4}{\sqrt[3]{19 - 3\sqrt{33}}} - \sqrt[3]{19 - 3\sqrt{33}}\right)$, and as $q \nearrow 1/2$, $p_q \searrow 0$. Thus, the regime over which a simple i.i.d. adversarial attack is optimal ($p \in [0, p_q]$) shrinks as the BSC gets noisier.
- For $q \in [0, 1/2)$, $C^{flip}(p, q)$, $p_q \leq p \leq 1/4$, is a decreasing linear function in p that intersects the p-axis at $p = 1/4$. Furthermore, $C^{flip}(p, q)$, $p_q \leq p \leq 1/4$, is in fact the tangent to the curve $1 - h_2(p \star q)$ at $p = p_q$. The optimal attack for Calvin in this regime relies on his snooping abilities and is based on a two-phase attack strategy described in section IV-B. The first phase of this attack involves Calvin injecting random i.i.d. bit-flips where the length of this phase is roughly $n\alpha(p, x)$. Therefore, (5) can be interpreted as an optimization over the lengths of the two attack phases. The analysis in Appendix B implies that the minimizer x^* in (5) is such that
$$\alpha(p, x^*) = \begin{cases} 1, & 0 \leq p \leq p_q \\ \frac{1-4p}{1-4p_q}, & p_q \leq p \leq 1/4 \end{cases}.$$
This corresponds to our earlier comment that for $p \in [0, p_q]$, $\alpha(p, x^*) = 1$, and it is optimal for Calvin to inject random i.i.d. noise across the entire codeword.

IV. Converse Proofs

To prove the converse, we demonstrate an attack strategy for Calvin in each of our models under which no rate larger than the claimed capacity expression is achievable. These attacks are inspired by, but different from, the attacks in [16], [17], [28] which only work when the erasure or the bit-flip probability is zero, i.e., $q = 0$. Specifically, our modified attacks rely crucially on Calvin's ability to snoop. We shall denote the transmitted and the received codewords as $\mathbf{x}$ and $\mathbf{y}$, respectively. The (stochastic) encoder and the decoder being used by Alice and Bob are denoted as $\tilde{\Phi}(\cdot|\cdot)$ and $\Gamma(\cdot)$, i.e. transmission rule $\mathcal{C} = (\tilde{\Phi}, \Gamma)$. Let $\mathbf{x}_L = (x_1, x_2, \cdots, x_\ell)$ and

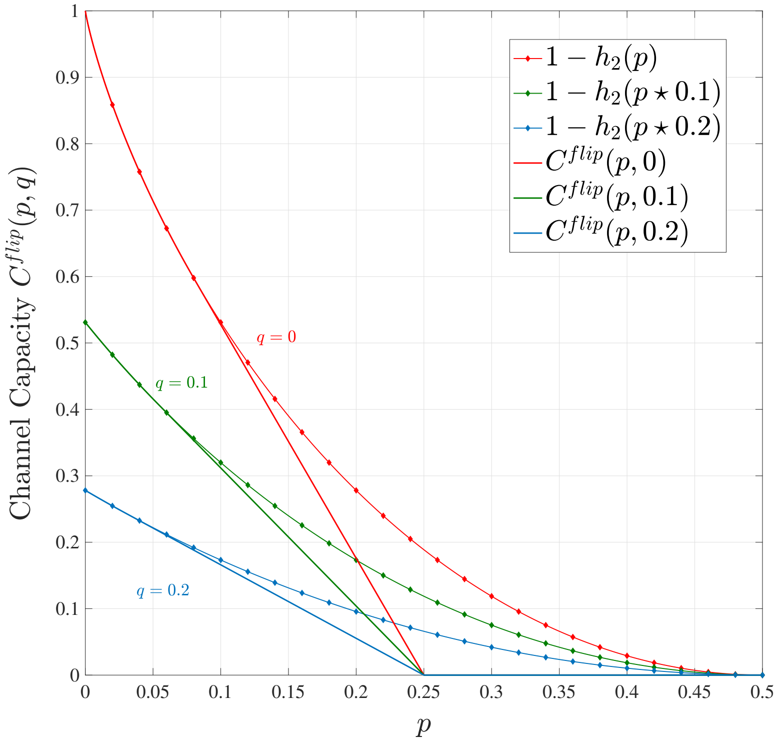


Fig. 2. The capacity $C^{flip}(p,q)$ of BSC(q)-ADV(p)-FS as a function of p. The cut-off value of p beyond which $C^{flip}(p,q)=0$ is $p=1/4$, independent of q.

$\mathbf{x}_R=(x_{\ell+1},\cdots,x_n)$, where ℓ is chosen later for each model. Similarly, let $\mathbf{y}_L=(y_1,y_2,\cdots,y_\ell)$ and $\mathbf{y}_R=(y_{\ell+1},\cdots,y_n)$.

A. *Converse for BEC(q)-ADV(p)-FS*

Our argument is based on a *wait and snoop, then push* attack. Suppose Alice attempts to communicate at a rate $R=C^{erase}(p,q)+\epsilon=(1-2p)(1-q)+\epsilon$. We will show that for sufficiently large block-length n, the probability of decoding error under the proposed attack is lower bounded by a constant that is only a function of ϵ (and independent of n). The attack constitutes of the following two phases:

- **Wait and snoop**: In this phase, Calvin waits and does not induce any erasures for the first $\ell=n\left(\frac{R-\frac{\epsilon}{2}}{1-q}\right)$ channel uses. Instead, Calvin simply snoops into Bob's reception to determine the erased/unerased bits and their positions. At the end of this phase, Bob receives $\mathbf{y}_L=(y_1,y_2,\cdots,y_\ell)$ containing some erased and some unerased bits. Note that the erasures

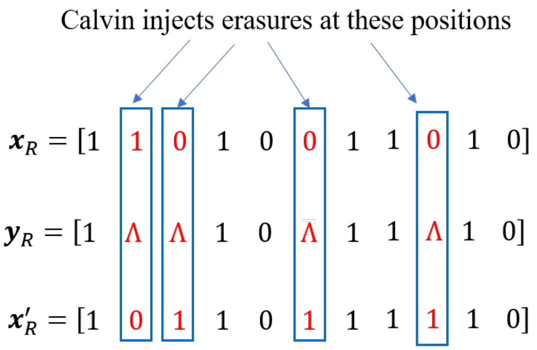

Calvin injects erasures at these positions
$\boldsymbol{x}_R = [1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0]$
$\boldsymbol{y}_R = [1\ \ \Lambda\ \ \Lambda\ \ 1\ \ 0\ \ \Lambda\ \ 1\ \ 1\ \ \Lambda\ \ 1\ \ 0]$
$\boldsymbol{x}_R' = [1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0]$

Fig. 3. In the push phase, if $\mathbf{x}_R$ and $\mathbf{x}_R'$ are sufficiently close (within distance pn), Calvin can make Bob completely uncertain whether the transmitted codeword was $\mathbf{x}$ or $\mathbf{x}'$.

in this phase occur purely due to the BEC(q) channel. Let $\{i_k\}_{k=1}^{m}$ be the indices of symbols in $\mathbf{y}_L$ that remain unerased.

- **Push**: Calvin forms the set $\mathcal{B}_{\mathbf{y}_L}$ of codewords consistent with $\mathbf{y}_L$ as

$$\mathcal{B}_{\mathbf{y}_L} = \{\mathbf{v} \in \mathcal{X}^n : \exists \tilde{u} \in \mathcal{U} \text{ s.t. } \tilde{\Phi}(\mathbf{v}|\tilde{u}) > 0 \text{ and } v_{i_k} = x_{i_k}\ k = 1, 2, \cdots, m\}, \tag{7}$$

where, as before, $\tilde{\Phi}(.|u)$ is the distribution of codewords selected when message u is to be transmitted. In other words, $\mathcal{B}_{\mathbf{y}_L}$ consists of all possible codewords that align with $\mathbf{y}_L$ at the positions that are unerased. Calvin then samples a codeword $\mathbf{x}'$ from $\mathcal{B}_{\mathbf{y}_L}$ according to the distribution $\mathbf{x}' \sim P_{\mathbf{X}|\mathbf{Y}_L=\mathbf{y}_L}(.|\mathbf{y}_L)$. In the push phase then, Calvin simply erases bit x_i, $i = \ell+1, \ell+2, \cdots n$ whenever $x_i \neq x_i'$, until his budget of pn erasures runs out.

During the push phase, if codewords $\mathbf{x}$ and $\mathbf{x}'$ correspond to distinct messages u and u' and we have $d(\mathbf{x}_R, \mathbf{x}_R') < pn$, then there would be no way for Bob to distinguish between messages u and u' and a decoding error would occur with probability at least $1/2$. This is illustrated in Fig. 3. We shall argue that this indeed occurs with a positive probability independent of n to settle the converse.

Analysis: The analysis is inspired from [17] where we also account for the presence of the BEC(q) in our claims. Define the set $A_0 = \left\{\mathbf{y}_L : H(\mathbf{U} \mid \mathbf{Y}_L = \mathbf{y}_L) > \frac{n\epsilon}{4}\right\}$ and the event $E_1 = \{\mathbf{Y}_L \in A_0\}$. We have the following lemma.

Lemma 5. $P(E_1) \geq \frac{\epsilon}{4}$.

Proof. Since $\mathbf{U} \to \mathbf{X}_L \to \mathbf{Y}_L$ is a Markov chain, by the data processing inequality, we have $I(\mathbf{U}; \mathbf{Y}_L) \leq I(\mathbf{X}_L, \mathbf{Y}_L) = \ell(1-q) = n(R - \epsilon/2)$. This holds since Calvin adds no erasures in the wait and snoop phase and the channel between $\mathbf{X}_L$ and $\mathbf{Y}_L$ is a BEC(q). Now, since $H(\mathbf{U}) = nR$, we have $H(\mathbf{U}|\mathbf{Y}_L) = \mathbb{E}_{\mathbf{Y}_L} H(\mathbf{U}|\mathbf{Y}_L = \mathbf{y}_L) = H(\mathbf{U}) - I(\mathbf{U};\mathbf{Y}_L) \geq n\epsilon/2$.. By Markov's inequality then, $P\left(nR - H(\mathbf{U}|\mathbf{Y}_L = \mathbf{y}_L) > nR - n\epsilon/4\right) \leq 1 - \frac{\epsilon/4}{R-\epsilon/4}$ which gives, $P(E_1) = P\left(H(\mathbf{U} \mid \mathbf{Y}_L = \mathbf{y}_L) > \frac{n\epsilon}{4}\right) \geq \frac{\epsilon}{4}$ as desired. □

Now let E_2 be the event $\{\mathbf{U} \neq \mathbf{U}'\}$ and E_3 be the event $\{d(\mathbf{X}_R, \mathbf{X}'_R) < pn\}$. First, we show the following.

Lemma 6. *For $\mathbf{y}_L \in A_0$, $P(E_2, E_3 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \epsilon^{\mathcal{O}(1/\epsilon)}$.*

Proof. Consider sampling $t = \frac{9}{\epsilon}$ codewords $\mathcal{C}_t = \left\{\mathbf{X}^{(1)}, \cdots, \mathbf{X}^{(t)}\right\}$ from the set $\mathcal{B}_{\mathbf{y}_L}$ where each codeword is sampled independently according to the conditional distribution $P_{\mathbf{X}|\mathbf{Y}_L=\mathbf{y}_L}(.|\mathbf{y}_L)$. Let the messages corresponding to the codewords be $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_t$ and let E_4 be the event that $\{\mathbf{U}_1, \mathbf{U}_2, \cdots \mathbf{U}_t$ are all distinct$\}$. We have from [17, A.2, Proposition 1] that for $\mathbf{y}_L \in A_0$ and for sufficiently large block length n,

$$P(E_4 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \left(\frac{\epsilon}{5}\right)^{t-1}. \tag{8}$$

Now, the average Hamming distance between the suffixes of codewords in $\mathcal{C}_t$ is defined as

$$d_{avg}(\mathcal{C}_t) = \frac{1}{t(t-1)} \sum_{i \neq j} d_H\left(\mathbf{X}_R^{(i)}, \mathbf{X}_R^{(j)}\right). \tag{9}$$

Conditioning on E_4, Plotkin's bound dictates

$$d_{avg}(\mathcal{C}_t) \leq \frac{1}{2}\frac{t}{t-1}(n-\ell) = n\frac{t}{t-1}\left(p - \frac{\epsilon}{4(1-q)}\right) \leq n\frac{\frac{9}{\epsilon}}{\frac{9}{\epsilon}-1}\left(p - \frac{\epsilon}{4}\right) \leq np - n\frac{\epsilon}{8}.$$

Thus for $\mathbf{y}_L \in A_0$, $\mathbb{E}(d_{avg}(\mathcal{C}_t) \mid E_4, \mathbf{Y}_L = \mathbf{y}_L) \leq np - n\epsilon/8$. Now, since all of the $\mathbf{X}^{(i)}$'s are picked independently, all pairs $(\mathbf{X}^{(i)}, \mathbf{X}^{(j)})$ have identical distribution. Thus,

$$\mathbb{E}(d_{avg}(\mathcal{C}_t) \mid E_4, \mathbf{Y}_L = \mathbf{y}_L) = \mathbb{E}(d_H(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) \mid E_4, \mathbf{Y}_L = \mathbf{y}_L) = \mathbb{E}(d_H(\mathbf{X}_R, \mathbf{X}'_R) \mid E_4, \mathbf{Y}_L = \mathbf{y}_L).$$

By Markov's inequality

$$P(d_H(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) > np \mid E_4, \mathbf{Y}_L = \mathbf{y}_L) \leq 1 - \frac{\epsilon}{8p}. \tag{10}$$

We have also that $P(E_2, E_3 \mid \mathbf{Y}_L = \mathbf{y}_L) = P(d(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) \leq pn, \mathbf{U}_1 \neq \mathbf{U}_2 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq P(d(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) \leq pn, E_4 \mid \mathbf{Y}_L = \mathbf{y}_L)$, where the last inequality holds because event E_4 is a subset of the event $\{U_1 \neq U_2\}$. Finally, from (8) and (10), we conclude $P(E_2, E_3 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \frac{\epsilon}{8p}\left(\frac{\epsilon}{5}\right)^{\frac{9}{\epsilon}-1} = \epsilon^{\mathcal{O}(1/\epsilon)}$. □

Recall that E_2 is the event that the message $\mathbf{U}'$ picked by the adversary is different from the one transmitted and E_3 is the event that the corresponding codewords $\mathbf{X}_R$ and $\mathbf{X}'_R$ are close enough so that Calvin's push phase succeeds and Bob is completely uncertain whether the message transmitted was $\mathbf{U}$ or $\mathbf{U}'$. Hence when E_2 and E_3 occur, the probability of decoding error is at least $1/2$. To finish the proof, we need only show a lower bound on $P(E_2, E_3)$. We have indeed, $P(E_2, E_3) \geq \sum_{\mathbf{y}_L \in A_0} P(E_2, E_3 \mid \mathbf{Y}_L = \mathbf{y}_L) P(\mathbf{Y}_L = \mathbf{y}_L) \geq \frac{\epsilon}{4}\frac{\epsilon}{8p}\left(\frac{\epsilon}{5}\right)^{\frac{9}{\epsilon}-1}$, a lower bound that is independent of n, hence settling the converse. We end this section with a few additional observations:

- *Feedback snooping helps*: After the wait and snoop phase, even though Calvin knows the entire prefix of the transmitted codeword $\mathbf{x}_L = (x_1, x_2, \cdots, x_\ell)$, he forms his set $\mathcal{B}_{\mathbf{y}_L}$ in (7) based only on the unerased bits. Intuitively, thanks to feedback snooping, Calvin exploits the additional equivocation induced by the BEC(q) to pick a random codeword from a larger set $\mathcal{B}_{\mathbf{y}_L}$, so that this codeword with high probability is sufficiently close to the transmitted codeword, and corresponds to a message different from the one that Alice chose.
- While we give Calvin full causal access to Bob's reception, an alternate model where Calvin is allowed *one-time block feedback* is sufficient - he would add no erasures for ℓ channel uses, retrieve through feedback the entire block $\mathbf{y}_L$ and then 'push'. Also note that interestingly, while the presence of the BEC(q) lowers the target rate, Calvin adds no erasures for approximately $n(1-2p)$ channel uses which from [15], [17] is also optimal when there is no BEC(q).
- Suppose Bob had access to an oracle who for each x_k that is erased informs him whether the erasure was due to Calvin, or the BEC(q), or both. It is straightforward to see that our converse proof continues to hold in this case. Thus, knowing *who* caused an erasure does not help Bob and the capacity is unchanged.

B. Converse for BSC(q)-ADV(p)-FS

Fix a $x \in [0, p]$. Suppose that for some $\epsilon > 0$, the transmitter attempts to communicate at a rate of $R = \alpha(p,x)\left(1 - h_2\left(\frac{x}{\alpha(p,x)} \star q\right)\right) + \epsilon$. We show that for sufficiently large n, under the proposed attack strategy for Calvin, the probability of decoding error in (1) is lower bounded by $\epsilon^{O(1/\epsilon)}$, a quantity *independent* of n. Since the same argument works for any x, the converse in theorem 2 holds. Our proof is based on a *babble and snoop, then push* attack that consists of the following two phases:

- **Babble and snoop**: For the first $\ell = (\alpha(p,x)+\epsilon/2)n$ channel uses, Calvin injects random bit-flips and monitors Bob's reception - at channel use i, $1 \leq i \leq \ell$, he flips bit x_i with probability xn/ℓ. At the end of this phase, Calvin knows $\mathbf{x}_L$ and $\mathbf{y}_L$.
- **Push**: Calvin samples a codeword $\mathbf{x}'$ (corresponding to message u') according to the conditional distribution $P_{\mathbf{X}|\mathbf{Y}_L=\mathbf{y}_L}(.|\mathbf{y}_L)$. His goal is to confuse the receiver between $\mathbf{x}$ and $\mathbf{x}'$. At positions where $\mathbf{x}_R$ and $\mathbf{x}'_R$ agree, he does nothing. Positions j where $\mathbf{x}_R$ and $\mathbf{x}'_R$ disagree, he flips x_j with probability $1/2$. This is illustrated in Fig. 4. This way, the Bob cannot distinguish between $\mathbf{x}$ and $\mathbf{x}'$ (even with the BSC(q)) due to the fact that $P(\mathbf{y}_R|\mathbf{x}_R) = P(\mathbf{y}_R|\mathbf{x}'_R)$. The proof relies on showing that with a small probability independent of n, u, u' are distinct and $\mathbf{x}_R$, $\mathbf{x}'_R$ are sufficiently close.

Calvin's attack requires knowledge of $\mathbf{Y}_L$, i.e., the symbols received by Bob during the first phase of the attack. Just like in the erasure model, the presence of the BSC(q) introduces additional equivocation at the receiver which Calvin is able to exploit to cause a reduction in rate. Here also, *one-time block feedback* (of entire block $\mathbf{y}_L$) after the first ℓ channel uses is sufficient for the attack to succeed.

Analysis: In the babble and snoop phase, by the Chernoff bound, Calvin uses at most $xn + \epsilon n/64$ flips with probability at least $1 - e^{-\Omega(\epsilon^2 n)}$. Let this be denoted as event E_1. Conditioned on E_1, Calvin's remaining budget in the push phase is atleast $(p-x)n - \epsilon n/64$. Define the set $A_0 = \left\{\mathbf{y}_L : H(\mathbf{U} \mid \mathbf{Y}_L = \mathbf{y}_L) > \frac{n\epsilon}{4}\right\}$. Denoting the event $E_2 = \{\mathbf{Y}_L \in A_0\}$, we have the following lemma.

Lemma 7. $P(E_2) \geq \epsilon/4$.

Proof. The proof is similar to claim 4 in [16]. $\mathbf{U} \to \mathbf{X}_L \to \mathbf{Y}_L$ is a markov chain and hence, by the data processing inequality and Calvin's actions in the babble phase, $I(\mathbf{U};\mathbf{Y}_L) \leq I(\mathbf{X}_L;\mathbf{Y}_L) = \ell\left(1 - h_2\left(\frac{xn}{\ell} \star q\right)\right)$. This is because the channel between $\mathbf{X}_L$ and $\mathbf{Y}_L$ is a cascade of $BSC(xn/\ell)$ and $BSC(q)$. Noting that $\ell = (\alpha + \epsilon/2)n$, we have $I(\mathbf{U};\mathbf{Y}_L) \leq n(\alpha + \epsilon/2)\left(1 - h_2\left(\frac{x}{\alpha+\epsilon/2} \star q\right)\right)$. Since $I(\mathbf{U},\mathbf{Y}_L) = H(\mathbf{U}) - H(\mathbf{U}|\mathbf{Y}_L)$ and $H(\mathbf{U}) = nR = n\alpha\Big(1 - h_2\left(\frac{x}{\alpha} \star q\right)\Big) + n\epsilon$, we get, $H(\mathbf{U}|\mathbf{Y}_L) \geq \frac{n\epsilon}{2} + n\left((\alpha+\epsilon/2)h_2\left(\frac{x}{\alpha+\epsilon/2} \star q\right) - \alpha h_2\left(\frac{x}{\alpha} \star q\right)\right)$. Now, the function $f(x) = xh_2\left(\frac{x}{x} \star q\right)$ is increasing in x, for any fixed $q \in (0,1/2)$. Hence, we have $H(\mathbf{U}|\mathbf{Y}_L) = \mathbb{E}_{\mathbf{Y}_L} H(\mathbf{U}|\mathbf{Y}_L = \mathbf{y}_L) \geq n\epsilon/2$. Finally, by Markov's inequality, $P(nR - H(\mathbf{U}|\mathbf{Y}_L = \mathbf{y}_L) > nR - n\epsilon/4) \leq 1 - \frac{\epsilon/4}{R-\epsilon/4}$ which gives, $P\left(H(\mathbf{U} \mid \mathbf{Y}_L = \mathbf{y}_L) > \frac{n\epsilon}{4}\right) \geq \frac{\epsilon}{4}$. □

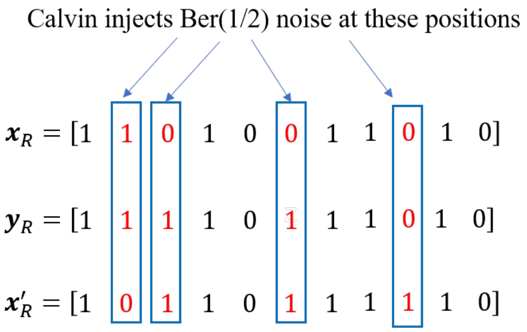

Calvin injects Ber(1/2) noise at these positions
$\boldsymbol{x}_R = [1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0]$
$\boldsymbol{y}_R = [1\ 1\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0]$
$\boldsymbol{x}'_R = [1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 0]$

Fig. 4. In the push phase, if $\mathbf{x}_R$ and $\mathbf{x}'_R$ are sufficiently close, Calvin can make Bob completely uncertain whether the transmitted codeword was $\mathbf{x}$ or $\mathbf{x}'$ by injecting Ber$(1/2)$ noise at positions where $\mathbf{x}_R$ differs from $\mathbf{x}'_R$.

Next, define the events $E_3 = \{\mathbf{U} \neq \mathbf{U}'\}$ and $E_4 = \{d_H(\mathbf{X}_R, \mathbf{X}'_R) \leq 2(p-x)n - \epsilon n/8\}$. E_3 is the event that the message picked by the adversary to confuse Bob in the push phase is different from the one transmitted. Similarly, event E_4 ensures that Calvin's remaining flips are enough to carry out his push attack. Using techniques from section A.2 of [17] and claim 6 in [16], we can now show the following.

Lemma 8. *For* $\mathbf{y}_L \in A_0$, $P(E_3, E_4 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \frac{\epsilon}{48}\left(\frac{\epsilon}{5}\right)^{\frac{12}{\epsilon}-1} = \epsilon^{\mathcal{O}(1/\epsilon)}$.

Proof. Consider sampling $t = \frac{12}{\epsilon}$ codewords $\mathcal{C}_t = \left\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(t)}\right\}$, each codeword sampled according to the conditional distribution $P_{\mathbf{X}|\mathbf{Y}_L=\mathbf{y}_L}(.|\mathbf{y}_L)$. Let the messages corresponding to the codewords be $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_t$ and let E_5 be the event that $\{\mathbf{U}_1, \mathbf{U}_2, \cdots \mathbf{U}_t$ are all distinct$\}$ i.e. all of the codewords are distinct. We have from proposition 1, section A.2 from [17] that for $\mathbf{y}_L \in A_0$, for sufficiently large block length n,

$$P(E_5 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \left(\frac{\epsilon}{5}\right)^{t-1}. \tag{11}$$

Recall that $\ell = (1 - 4(p-x) + \epsilon/2)n$. Conditioning on E_5, by Plotkin's bound we have $d_{avg}(\mathcal{C}_t) \leq \frac{1}{2}\frac{t}{t-1}(n-\ell) \leq 2(p-x)n - \epsilon n/6$, where $d_{avg}(\mathcal{C}_t)$ is defined in (9). Thus for $\mathbf{y}_L \in A_0$, $\mathbb{E}(d_{avg}(\mathcal{C}_t) \mid E_5, \mathbf{Y}_L = \mathbf{y}_L) \leq 2(p-x)n - \epsilon n/6$. Now, since all of the $\mathbf{X}^{(i)}$'s are picked independently, all pairs $(\mathbf{X}^{(i)}, \mathbf{X}^{(j)})$ have identical distribution. Thus, $\mathbb{E}(d_{avg}(\mathcal{C}_t) \mid E_5, \mathbf{Y}_L = \mathbf{y}_L) = \mathbb{E}(d_H(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) \mid E_5, \mathbf{Y}_L = \mathbf{y}_L) = \mathbb{E}(d_H(\mathbf{X}_R, \mathbf{X}'_R) \mid E_5, \mathbf{Y}_L = \mathbf{y}_L)$. By Markov's inequality,

$$P(d_H(\mathbf{X}_R^{(1)}, \mathbf{X}_R^{(2)}) > 2(p-x)n - \epsilon n/8 \mid E_5, \mathbf{Y}_L = \mathbf{y}_L) \leq \frac{2(p-x)n - \epsilon n/6}{2(p-x)n - \epsilon n/8} \leq 1 - \frac{\epsilon}{48}. \tag{12}$$

Following the arguments as in the proof of Lemma 6, (11) and (12) imply that for $\mathbf{y}_L \in A_0$, $P(E_3, E_4 \mid \mathbf{Y}_L = \mathbf{y}_L) \geq \frac{\epsilon}{48}\left(\frac{\epsilon}{5}\right)^{\frac{12}{\epsilon}-1}$. $\square$

Now, in the push phase, Calvin injects $Ber(1/2)$ noise at $d_H(\mathbf{X}_R, \mathbf{X}_R^{'})$ positions. Conditioned on E_1, Calvin has at least a budget of $(p-x)n - \epsilon n/64$ bit-flips that remain. If $\mathbf{a}_R$ is the error vector chosen by Calvin in the push phase, conditioned on E_3 and E_4 we have $\mathbb{E}(d_H(\mathbf{a}_R, \mathbf{0})) = (p-x)n - \epsilon n/16$. Further by the Chernoff bound, with probability at least $1-2^{-\Omega(\epsilon^2 n)}$, the distance $d_H(\mathbf{a}_R, \mathbf{0})$ is within $3\epsilon n/64$ of its expected value. Let this event be E_5. Since $\mathbb{E}(d_H(\mathbf{a}_R, \mathbf{0})) + 3\epsilon n/64 = (p-x)n - \epsilon n/64$, the power constraint is respected w.h.p..

When events E_1, E_3, E_4, E_5 occur, the probability of decoding error is clearly at least $1/2$ since the receiver cannot distinguish between $\mathbf{x}$ and $\mathbf{x}'$. Since $P(E_1) \geq 1 - e^{-\Omega(\epsilon^2 n)}$ and $P(E_5) \geq 1 - e^{-\Omega(\epsilon^2 n)}$, the bound in Lemma 8 together with the bound $P(E_2) \geq \epsilon/4$ implies for sufficiently large n, the maximum probability of error in (1) is at least of the order $\epsilon^{O(1/\epsilon)}$, a quantity independent of n and the proof is complete.

V. Achievability Proofs

To prove achievability, we resort to a random coding argument. Unfortunately, the classical random deterministic code ensemble where the (unique) codeword for each message is drawn independently and uniformly randomly, does not work. A modification of Calvin's attacks from our converse proofs in section IV defeats such an attempt. Indeed, consider for instance the BEC(q)-ADV(p)-FS channel with $q = 0$. The claimed capacity expression is $C^{erase}(p, 0) = 1-2p$. Suppose Alice wants to transmit at rate $R = 1 - 2p - \epsilon$. Let Ψ denote the codebook containing 2^{nR} length-n codewords. We argue that the probability a randomly chosen Ψ enables reliable communication goes to 0 as $n \to \infty$. First, it can be shown that with probability approaching 1 as $n \to \infty$, randomly sampled Ψ satisfies the following : Ψ contains a codeword $\mathbf{x}$ for which

- At least $2^{n\epsilon/2}$ codewords other than $\mathbf{x}$ share the same prefix $(x_1, x_2, \cdots, x_{nR-n\epsilon})$.
- No other codeword has the prefix $(x_1, x_2, \cdots, x_{nR+n\epsilon})$.

For such a codeword, consider the following attack for Calvin:

- **Wait**: For the first $R - \epsilon$ channel uses, Calvin adds no erasures. Bob (and Calvin) narrow down the transmitted codeword to a list $\mathcal{L}$ which is of size at least $2^{n\epsilon/2}$.
- **Block**: For the next subsequent $2\epsilon n$ channel uses, Calvin erase all of the bits. Calvin who knows $(x_1, \cdots, x_{nR+n\epsilon})$ determines the transmitted codeword $\mathbf{x}$.

- **Push**: Calvin picks $\tilde{\mathbf{x}} \in \mathcal{L}, \tilde{\mathbf{x}} \neq \mathbf{x}$ that minimizes $d_H((x_{nR+n\epsilon+1}, \cdots, x_n), (\tilde{x}_{nR+n\epsilon+1}, \cdots, \tilde{x}_n)) = d_H(\mathbf{x}_P, \tilde{\mathbf{x}}_P)$ and injects an erasure at channel use j if $x_j \neq \tilde{x}_j$, until his budget runs out. Calvin succeeds in confusing Bob between $\mathbf{x}$ and $\tilde{\mathbf{x}}$ if $d_H(\mathbf{x}_P, \tilde{\mathbf{x}}_P) < pn - 2\epsilon n$.

The length of the push phase is $2pn$ and $|\mathcal{L}| \geq 2^{n\epsilon/2}$. It can be shown that with probability approaching 1 as $n \to \infty$, $2^{n\epsilon/2}$ codewords of length $2pn$ picked uniformly randomly have minimum distance less than $pn - 2\epsilon n$ [29]. Thus, the random deterministic code ensemble does not work. A similar argument can also be made for the BSC(q)-ADV(p)-FS channel.

Therefore, we will consider instead an ensemble of *stochastic* codes and show that with positive probability, a stochastic code drawn randomly from the ensemble enables reliable communication between Alice and Bob. For both channel models BEC(q)-ADV(p)-FS and BSC(q)-ADV(p)-FS, we shall use the code ensemble from [15] with reduced rates as given in theorems 1 and 2 respectively. However, note that compared to the $q = 0$ case, the decoding procedure and analysis will need to be modified greatly to deal with compounded adversarial and random errors.

Random code distribution: Alice is endowed with a set of private keys or secrets for encoding, $\mathcal{S} = \{1, 2, \cdots, 2^{nS}\}$. The encoding procedure is carried out in chunks, each of size $n\theta$ where $\theta < 1$ is a quantization parameter. The values for S and θ are set specific to the coding rate and the channel model during analysis later. Let Ξ be the uniform distribution over stochastic codes $\mathcal{C} : \mathcal{U} \times \mathcal{S} \to \mathcal{X}^{n\theta}$, i.e., for each $u \in \mathcal{U}$ and $s \in \mathcal{S}$, $\mathcal{C}(u, s)$ is picked independently and uniformly randomly. Then each chunk i, $1 \leq i \leq \frac{1}{\theta}$, is associated to a stochastic code $\mathcal{C}_i$ drawn independently from the distribution Ξ. The transmission rule is composed of maps $\mathcal{C}_1, \mathcal{C}_2, \cdots \mathcal{C}_{1/\theta}$ and a decoder.

Encoding procedure: For message $u \in \mathcal{U}$ and keys $s_1, s_2, \cdots, s_{\frac{1}{\theta}}$, the codeword $\mathbf{x}$ selected for transmission is $\mathbf{x} = \mathcal{C}_1(u, s_1) \circ \mathcal{C}_2(u, s_2) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u, s_{\frac{1}{\theta}})$, where $\circ$ represents the concatenation operator. We refer to $\mathcal{C}_i(u, s_i)$ as the i^{th} sub-codeword or the i^{th} chunk and the code $\mathcal{C}_i$ as the i^{th} sub-code. Each secret or key s_i for encoding with $\mathcal{C}_i$ is chosen uniformly randomly from $\mathcal{S}$.

Decoding: The decoding is specific to each of the channel models and described shortly.

Definitions: Define the set $\mathcal{T} = \{n\theta, 2n\theta, \cdots, n - n\theta\}$ containing indices of the chunk ends. For some $t \in \mathcal{T}$ where $t = kn\theta$, we refer to $\mathcal{C}_1 \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_k$ as the left mega sub-code w.r.t. t and $\mathcal{C}_{k+1} \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_{\frac{1}{\theta}}$ as the right mega sub-code w.r.t. t. Accordingly, the concatenation of the first k sub-codewords is be referred to as the left mega sub-codeword w.r.t. t, and that of the last $\frac{1}{\theta} - k$ sub-codewords is referred to as right mega sub-codeword w.r.t. t. We shall also denote the key sequences used to encode the left and the right mega-subcodewords as $s_{left} = (s_1, s_2, \cdots, s_k)$

and $s_{right} = (s_{k+1}, s_{k+2}, \cdots, s_{\frac{1}{\theta}})$.

A. Achievability for BEC(q)-ADV(p)-FS

Fix $\epsilon' > 0$ and let $R = (1-2p)(1-q) - \epsilon' = (1-2p-\epsilon)(1-q)$, where $\epsilon = \epsilon'/(1-q)$. We show that R is achievable for the BEC(q)-ADV(p)-FS model. We set initially $\theta = \frac{\epsilon}{4}$ and $S = \frac{\theta^3}{8}$.

Decoding procedure: A clean codeword $\mathbf{x}$ is said to be consistent with corrupted word $\mathbf{y}$ if $\mathbf{x}$ and $\mathbf{y}$ agree on the unerased positions. The decoding procedure for the BEC(q)-ADV(p)-FS channel is very simple. Bob decodes the received word $\mathbf{y}$ to the unique message $\hat{u}$ for which there is at least one associated codeword that is consistent with $\mathbf{y}$. If more than one such message exists, a decoding error is declared. Mathematically, Bob forms the list of consistent messages

$$\mathcal{L} = \{u \in \mathcal{U} : \exists\ (s_1, \cdots, s_{1/\theta}) \in \mathcal{S}^{1/\theta} \text{ s.t. } \mathcal{C}_1(u, s_1) \circ \cdots \mathcal{C}_{1/\theta}(u, s_{1/\theta}) \text{ and } \mathbf{y} \text{ are consistent}\},$$

and decodes successfully when $\mathcal{L}$ has exactly one single message.

Analysis: For the analysis, we work with an alternate two-phase but equivalent view of the decoding process. This also allows us to give a unified view of decoding for both channel models. For some $t^* = k^* n\theta$, partition received word $\mathbf{y}$ into $\mathbf{y}_1^{t^*} = (y_1, \ldots, y_{t^*})$ and $\mathbf{y}_{t^*+1}^n = (y_{t^*+1}, \ldots, n)$. Decoding can be split into two sequential phases.

- <u>List decoding</u>: Perform list decoding on $\mathbf{y}_1^{t^*}$ to obtain the list of messages $\mathcal{L}$ that are consistent with Bob's reception $\mathbf{y}_1^{t^*}$.

$$\mathcal{L} = \{u \in \mathcal{U} : \exists\ (s_1, \cdots, s_{k^*}) \in \mathcal{S}^{k^*} \text{ s.t. } \mathcal{C}_1(u, s_1) \circ \cdots \mathcal{C}_k(u, s_{k^*}) \text{ and } \mathbf{y}_1^{t^*} \text{agree}\}.$$

- <u>Unique decoding or list refinement</u>: Refine the list by removing all messages in $\mathcal{L}$ that are not consistent with $\mathbf{y}_{t^*+1}^n$.

$$\mathcal{L}^{ref} = \{u \in \mathcal{L} : \exists\ (s_{k^*+1}, \cdots, s_{1/\theta}) \in \mathcal{S}^{1/\theta - k^*} \text{ s.t. } \mathcal{C}_{k^*+1}(u, s_{k^*+1}) \circ \cdots \mathcal{C}_{1/\theta}(u, s_{1/\theta}) \text{ and } \mathbf{y}_{t^*+1}^n \text{ agree}\}.$$

If exactly one message, say $\hat{u}$, remains in $\mathcal{L}$ after refinement, the decoder outputs $\hat{u}$. If the refined list $\mathcal{L}^{ref}$ does not contain exactly one message, a decoding error is declared. Decoding is successful if $\hat{u} = u^*$, the true message transmitted by Alice.

The proof involves showing that there is a value of $t^* \in \mathcal{T}$ that Bob can choose for which decoding succeeds. When $q = 0$, as shown in [15], t^* is chosen as a function of the number of (purely adversarial) erasures $\lambda^a_{t^*}$ observed in $\mathbf{y}$ up until time t^*. Specifically, Bob chooses t^* as the smallest integer that satisfies the so-called *list-decoding condition*

$$\lambda^a_{t^*} \leq t^*(1-\theta) - ((1-2p)-\epsilon)n \tag{13}$$

and the *energy bounding condition*

$$np - \lambda^a_{t^*} \leq \frac{(n-t^*)(1-\theta)}{2}. \tag{14}$$

Condition (13) ensures the size of $\mathcal{L}$ is small (at most a constant) while condition (14) ensures the fraction of erasures that occur in $\mathbf{y}^n_{t^*+1}$ is small enough to perform list refinement. When $q > 0$, we modify conditions (13) and (14) appropriately for the BEC(q)-ADV(p)-FS channel.

Choice of t^* when $q > 0$: Let λ^a_t be the number of erasures injected adversarially by Calvin up until t and let λ_t denote the number of erasures observed by Bob up until time t, which includes contributions both from Calvin and the BEC(q). Bob chooses t^* as the smallest integer that satisfies simultaneously the modified list-decoding condition

$$\lambda_{t^*} - qt^* \leq t^*(1-q)(1-\theta) - Rn \tag{15}$$

and the modified list refinement condition

$$np(1-q) - (\lambda_{t^*} - qt^*) \leq \frac{(n-t^*)(1-q)(1-\theta)}{2}. \tag{16}$$

From Lemma 3, if Calvin adds $\lambda^a_{t^*}$ erasures up until t^*, the total number of erasures λ_{t^*} that Bob observes is approximately $\lambda_{t^*} \approx \lambda^a_{t^*} + q(t^* - \lambda^a_{t^*})$. On making this substitution we see that t^* satisfying (15) and (16) is nearly the same as that satisfying (13) and (14) i.e. it is sufficient to choose t^* only as a function of purely adversarial erasures. However, since Bob has no way of knowing this, he works with the quantity $\lambda_{t^*} - qt^*$. Note that since qt^* is an estimate of the number of erasures added by the BEC(q), we can interpret $\lambda_{t^*} - qt^*$ to be an estimate of the number of adversarial erasures that *do not* coincide with random erasures.

Having selected t^*, Bob can then finish decoding using the two-phase decoding process described previously to successfully recover w.h.p. the transmitted message. We will now prove this. First, by Lemma 3, for $\delta > 0$, the total number of erasures that Bob observes at time $t = kn\theta$ satisfies $\lambda_t \in [\lambda^a_t + (t-\lambda^a_t)(q-\delta), \lambda^a_t + (t-\lambda^a_t)(q+\delta)]$ with probability at least $1 - 2^{-\Omega(\delta^2 n)}$. Thus, Bob's choice t^* satisfies w.h.p.

$$\lambda_{t^*} - qt^* = \hat{\lambda}_{t^*} \in [\lambda^a_{t^*}(1-q+\delta) - \delta t^*, \lambda^a_{t^*}(1-q-\delta) + \delta t^*]. \tag{17}$$

Let $\mathcal{Z} = [\lambda^a_{t^*}(1-q+\delta) - \delta t^*, \lambda^a_{t^*}(1-q-\delta) + \delta t^*]$. By a similar analysis as in [15, Claim B.3], we show in Lemma 9 that when $\delta > 0$ is small enough, a $t^* \in \mathcal{T}$ exists that satisfies both (15) and (16), for any realization of $\hat{\lambda}_{t^*} \in \mathcal{Z}$. The proof of Lemma 9 can be found in Appendix A.

Lemma 9. *We can choose a $\delta > 0$ such that the following holds : for any strategy selected by Calvin, with probability at least $1 - 2^{-\Omega(\delta^2 n)}$, there exists a $t^* \in \mathcal{T}$ such that both of the following conditions hold:*

$$\lambda_{t^*} - qt^* \leq t^*(1-q)(1-\theta) - Rn, \text{ and } np(1-q) - (\lambda_{t^*} - qt^*) \leq \frac{(n-t^*)(1-q)(1-\theta)}{2}.$$

Calvin's unused budget: We now give an upper bound on the number of adversarial erasures that Calvin is left with to add on to the right mega sub-codeword. Since the total budget is pn, the remaining number of erasures is $pn - \lambda^a_{t^*}$. From (16) and (17), for any $\hat{\lambda}_{t^*} \in \mathcal{Z}$, we have $pn - \lambda^a_{t^*} \leq \frac{(n-t^*)(1-\theta)}{2} + \frac{\delta(t-\lambda^a_{t^*})}{1-q}$. Since we are proving an achievability result and θ is representative of the back-off from the capacity expression, we can choose θ as small as we would like. Choosing θ sufficiently small so that for instance $\delta = \frac{1}{4}\frac{(1-q)\theta^2(1-\theta)}{1+2\theta-\theta^2} \leq \frac{1}{16}(1-q)\theta^2$, we get the bound

$$pn - \lambda^a_{t^*} \leq (n-t^*)\left(\frac{1}{2} - \frac{7\theta}{16}\right). \tag{18}$$

List decoding: We show that with probability at least $\left(1-\frac{1}{n}\right)$ over the code design, the size of the list of messages $\mathcal{L}$ obtained by Bob in the list-decoding phase is at most a constant, specifically, $|\mathcal{L}| < C/\epsilon$ for some constant C.

Lemma 10. *(Modified from [15, Claims B.5-B.7]) Let $t^* \in \mathcal{T}$ where $t^* = k^* n\theta$. For sufficiently large n, with probability at least $\left(1-\frac{1}{n}\right)$ over the code design, the left mega sub-code $\mathcal{C}_1 \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_{k^*}$ is list decodable with list size $L = O\left(\frac{1}{\epsilon}\right)$ for λ_{t^*} erasures where t^* and λ_{t^*} satisfy* (15), *i.e., $\lambda_{t^*} - qt^* \leq t^*(1-q)(1-\theta) - Rn$.*

Proof. The proof follows exactly the analysis in [15, Claims B.5-B.7]. The only additional step is to verify if the bound $1 - \frac{\lambda_{t^*}}{t^*} - \frac{nR}{t^*} - \frac{S}{\theta} \geq \frac{\theta}{2}$ holds. This is indeed the case as we have

$$1 - \frac{\lambda_{t^*}}{t^*} - \frac{nR}{t^*} - \frac{S}{\theta} - \frac{\theta}{2} \overset{(a)}{=} \frac{1}{t^*}(1 - \lambda_{t^*} - nR) - \frac{\theta^2}{8} - \frac{\theta}{2} \overset{(b)}{\geq} \theta(1-q) - \frac{\theta^2}{8} - \frac{\theta}{2} \overset{(c)}{\geq} 0,$$

where (a) follows from the substitution $S = \frac{\theta^3}{8}$, (b) follows from (15) and (c) holds by choosing θ sufficiently small. □

List refinement: For some chunk end $t \in \mathcal{T}$ where $t = kn\theta$, $\mathbf{y}_1^t = (y_1, y_2, \cdots, y_t)$ and $\mathbf{y}_{t+1}^n = (y_{t+1}, \cdots, y_n)$ are the left mega received word and the right mega received word w.r.t. t respectively. Let u^* be the true message chosen by Alice for transmission. Given any list of messages $\mathcal{L}$, we define $\mathcal{L}(u^*)$ to be the set of all possible right mega sub-codewords w.r.t t for each message in $\mathcal{L} \setminus \{u^*\}$ i.e.

$$\mathcal{L}(u^*) = \{\mathcal{C}_{k+1}(u, s_{k+1}) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u, s_{\frac{1}{\theta}}) : u \in \mathcal{L}, u \neq u^*, (s_{k+1}, \cdots, s_{1/\theta}) \in \mathcal{S}^{\frac{1}{\theta}-k}\}.$$

For convenience, we enumerate $\mathcal{L}(u^*)$ containing codewords of length-$(n-t)$ as $\mathcal{L}(u^*) = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{|\mathcal{L}(u^*)|}\}$. The right mega-subcodeword for the true message is $\mathbf{x}_{t+1}^n(u^*, s_{right}) = \mathcal{C}_{k+1}(u^*, s_{k+1}) \circ \mathcal{C}_{k+2}(u^*, s_{k+2}) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u^*, s_{\frac{1}{\theta}})$, which we emphasize is a function of the specific realization of $s_{right=}(s_{k+1}, \cdots, s_{\frac{1}{\theta}})$ during encoding. For any list such that $|\mathcal{L}|\leq O\left(\frac{1}{\epsilon}\right)$, we would like our code design to satisfy the following distance condition

$$d_H\left(\mathbf{x}_{t+1}^n(u^*, s_{right}), \mathbf{w}_j\right) \geq (n-t)\left(\frac{1}{2} - \frac{3\theta}{8}\right) \quad \forall \mathbf{w}_j \in \mathcal{L}(u^*). \tag{19}$$

Equation (19) is a key property that guarantees successful decoding. It ensures that the right mega sub-codeword for the transmitted message is sufficiently far in Hamming distance from the right mega sub words for any of the other messages in list $\mathcal{L}$ that is obtained by Bob during list decoding. We show that (19) indeed occurs w.h.p., for almost all possible sequence of secrets s_{right}.

Lemma 11. *(Modified from [15, Claims B.11-B.14]) Let $C > 0$ be an arbitrary constant. Then, there is a n_0 such that for $n > n_0$, with probability at least $1 - 2^{-n}$, a code drawn from our random ensemble satisfies the following property : for every chunk end $t \in \mathcal{T}$, for every message u^*, and every list $\mathcal{L}$ of size at most C/ϵ, we have that* (19) *holds for at least a* $(1 - 2^{-nS/4})$ *portion of all possible secret sequences s_{right}.*

Proof. Fixing a sequence of secrets $s_{right} = (s_{k+1}, \cdots, s_{\frac{1}{\theta}})$, message u^* and list $\mathcal{L}$, we first show that (19) holds w.h.p.. Let radius $r = \left(\frac{1}{2} - \frac{3\theta}{8}\right)$. We surround each word $\mathbf{w}_j \in \mathcal{L}(u^*)$ with a Hamming ball of radius r and the union of all the balls is the so called forbidden region.

For (19) to hold, we must have that $\mathbf{x}_{t+1}^n(u^*, s_{right})$ is outside all these balls, i.e. outside the forbidden region. Due to the code construction, $\mathbf{x}_{t+1}^n(u^*, s_{right})$ is uniformly distributed over all possible binary vectors of length $(n-t)$ and thus it is enough to bound the size of the forbidden region. If the size of the list $\mathcal{L}$ is L, the size of $\mathcal{L}(u^*)$ is at most $L.2^{nS\left(\frac{1}{\theta}-\frac{t}{n\theta}\right)}$. Hence the number of codewords in the forbidden region is at most

$$L.2^{nS\left(\frac{1}{\theta}-\frac{t}{n\theta}\right)}\sum_{j=0}^{r}\binom{n-t}{j} < 2^{(n-t)\left(\frac{\log_2 L}{n-t}+\frac{S}{\theta}+h_2\left(\frac{1}{2}-\frac{3\theta}{8}\right)\right)}.$$

From the Taylor expansion of function $h_2(x)$ in a neighborhood of $1/2$, we can show $h_2\left(\frac{1}{2}-\frac{3\theta}{8}\right) < 1-\frac{9\theta^2}{32\ln(2)}$. Let $\eta = \frac{\theta^2}{4}$. For sufficiently large n, we have

$$\left(\frac{\log_2 L}{n-t}+\frac{S}{\theta}+h_2\left(\frac{1}{2}-\frac{3\theta}{8}\right)\right) < \left(\frac{\log_2 L}{n-t}+\frac{S}{\theta}+\left(1-\frac{9\theta^2}{32\ln(2)}\right)\right) < 1-\eta.$$

Hence, the total number of codewords in the forbidden region is at most $2^{(n-t)(1-\eta)}$ and we have

$$P\Big(\mathbf{x}_{t+1}^n(u^*, s_{right}) \text{ is outside the forbidden region}\Big) > \frac{2^{(n-t)}-2^{-(n-t)(1-\eta)}}{2^{n-t}} = 1-2^{-(n-t)\eta}.$$

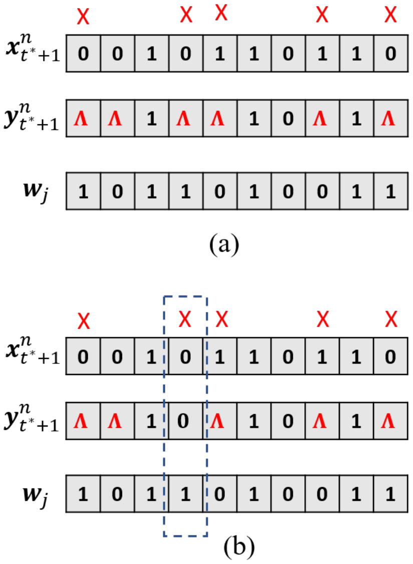
From here on, the rest of the steps in the proof follow claims B.12-B.14 in [15]. □

Success of decoding: From the preceding discussion, there exists a code in our random ensemble that satisfies the following simultaneously (irrespective of Calvin's strategy):

- For t^* selected according to (15) and (16), the size of the list $\mathcal{L}$ obtained by Bob during list decoding is at most C/ϵ for some constant C. Further, the transmitted message u^* is inside list $\mathcal{L}$.
- For almost all possible realizations of secret sequences s_{right} (at least a fraction $1-2^{-nS/4}$ of them), the right mega codeword corresponding to message u^* denoted $\mathbf{x}_{t^*+1}^n(u^*, s_{right})$, is at least $(n-t^*)\left(\frac{1}{2}-\frac{3\theta}{8}\right)$ away in Hamming distance from any codeword in $\mathcal{L}(u^*)$.

Recall from (18) that with probability at least $1-2^{-\Omega(\delta^2 n)}$, Calvin has at most $pn-\lambda_{t^*}^a \leq (n-t^*)\left(\frac{1}{2}-\frac{7\theta}{16}\right)$ erasures that remain. Consider any arbitrary codeword $\mathbf{w}_j \in \mathcal{L}(u^*)$ that is associated with message $u' \neq u^*$. Let $\mathcal{I}^c$ be the set of indices where $\mathbf{w}_j$ and $\mathbf{x}_{t^*+1}^n(u^*, s_{right})$ disagree. The only way that Bob is unable to distinguish between $\mathbf{w}_j$ and $\mathbf{x}_{t^*+1}^n(u^*, s_{right})$ and hence makes a decoding error of at least $1/2$ is when indices $\mathcal{I}^c$ in $\mathbf{y}_{t^*+1}^n$ are all erased due to Calvin and the BEC(q). In other words, if $\mathcal{J}$ is the set of indices of erasures in $\mathbf{y}_{t^*+1}^n$, we must have $\mathcal{J} \supset \mathcal{I}^c$. An example is illustrated in Fig. 5.

Now clearly, if Calvin wishes to confuse Bob between $\mathbf{w}_j$ and $\mathbf{x}_{t^*+1}^n(u^*, s_{right})$, his best strategy is to add all erasures at positions $\mathcal{I}^c$. However, this still leaves at least $(n-t^*)\left(\frac{1}{2}-\frac{3\theta}{8}\right)-(n-t^*)\left(\frac{1}{2}-\frac{7\theta}{16}\right) = (n-t^*)\frac{\theta}{16}$ positions where $\mathbf{w}_j$ and $\mathbf{x}_{t+1}^n(u^*, s_{right})$ disagree but no adversarial erasures are added. For Bob to be confused between $\mathbf{w}_j$ and $\mathbf{x}_{t+1}^n(u^*, s_{right})$, the BEC($q$) must erase *all of the* $(n-t^*)\frac{\theta}{16}$ bits that Calvin could not erase. However, this event occurs with


X X X X X
$x^n_{t^*+1}$ 0 0 1 0 1 1 0 1 1 0
$y^n_{t^*+1}$ Λ Λ 1 Λ Λ 1 0 Λ 1 Λ
w_j 1 0 1 1 0 1 0 0 1 1
X X X X X
$x^n_{t^*+1}$ 0 0 1 0 1 1 0 1 1 0
$y^n_{t^*+1}$ Λ Λ 1 0 Λ 1 0 Λ 1 Λ
w_j 1 0 1 1 0 1 0 0 1 1

(a)

(b)

Fig. 5. In (a), the set of indices in Bob's observation $\mathbf{y}^n_{t^*+1}$ where $\mathbf{x}^n_{t^*+1}$ and $\mathbf{w}_j$ differ are all erased. Therefore, Bob cannot determine if Alice transmitted $\mathbf{x}^n_{t^*+1}$ or $\mathbf{w}_j$. In (b), successful reception of even one bit where $\mathbf{x}^n_{t^*+1}$ and $\mathbf{w}_j$ disagree allows Bob to disambiguate between $\mathbf{x}^n_{t^*+1}$ and $\mathbf{w}_j$.

probability $q^{(n-t^*)\frac{\theta}{16}} \leq 2^{-n\Omega(\theta^2)}$. Thus, the probability of the error event that Bob cannot distinguish between $\mathbf{y}^n_{t^*+1}$ and $\mathbf{w}_j$ is exponentially small. Repeating the same argument for any $\mathbf{w}_j \in \mathcal{L}(u^*)$, we have that a decoding error occurs with exponentially small probability. Thus, Bob succeeds in determining the transmitted message u^* and the proof is complete.

B. Achievability for BSC(q)-ADV(p)-FS

Let $\epsilon > 0$ such that $p' = p + \frac{\epsilon^2}{16} < \frac{1}{4}$. We also set $\theta = \frac{\epsilon^2(1-4p')}{4}$, $S = \frac{\theta^3}{8}$. To show that $C^{flip}(p,q)$ in Theorem 2 is the capacity, we let the rate be $R = C^{flip}(p',q) - \epsilon$ and prove that for any $\delta > 0$ and every sufficiently large block length n, a randomly sampled stochastic code $\mathcal{C}$ with rate R satisfies $P_e(\mathcal{C}) < \delta$ with a positive probability.

Decoding procedure: Recall that the received codeword can be written as $\mathbf{y} = \mathbf{x} \oplus \mathbf{a} \oplus \mathbf{z}$ where $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ is the adversarial error vector added by Calvin and $\mathbf{z} = (z_1, z_2, \cdots, z_n)$ is the error vector produced by the BSC(q). In accordance to the power constraint, we have $d_H(\mathbf{a}, \mathbf{0}) \leq pn$. Positions i where $a_i = z_i = 1$, symbols x_i remain unflipped.

To describe the decoding process, we need to define certain quantities. For a chunk end $t \in \mathcal{T}$, let p_t be the normalized number of bit-flip attempts used up by Calvin up until time t i.e.

$$p_t = \frac{weight\{(a_1, a_2, \cdots, a_t)\}}{t}.$$

Note that since p_t only captures adversarial error injections, the word received up until time t may have more or less effective bit-flips than tp_t. For the purposes of decoding, Bob maintains a reference $\hat{p}_t$ which is approximately defined as follows: $\hat{p}_t = \frac{n}{t}\left(p - \frac{1}{4}\right) + \frac{1}{4}$ for $t \geq n(1-4p)$ and $\hat{p}_t = 0$ for $t < n(1-4p)$. It can be seen that $\hat{p}_t$ in increasing in $t \in [n(1-4p), n]$ reaching $\hat{p}_n = p$ as is expected. For a rigorous analysis, certain twiddle terms need to be added to this definition as is explained later. We shall refer to p_t as the true trajectory and $\hat{p}_t$ as the reference trajectory for adversarial bit-flip attempts.

The overall decoding process is iterative potentially involving several decoding attempts. For some chunk end $t \in \mathcal{T}$ where $t = kn\theta$, $\mathbf{y}_1^t = (y_1, y_2, \cdots, y_t)$ and $\mathbf{y}_{t+1}^n = (y_{t+1}, \cdots, y_n)$ are the left mega received word and the right mega received word w.r.t. t. Similarly, $\mathbf{x}_1^t = (x_1, x_2, \cdots, x_t)$ and $\mathbf{x}_{t+1}^n = (x_{t+1}, \cdots, x_n)$ are the left mega transmitted codeword and the right mega transmitted codeword w.r.t. t. A decoding attempt w.r.t t consists of two phases - a list-decoding phase followed by a unique decoding phase.

List decoding: In the list decoding phase, Bob identifies the set of messages for whom there is at least one associated codeword whose left mega sub-codeword w.r.t. t is within Hamming distance $t(\hat{p}_t \star q + \delta_1)$ from $\mathbf{y}_1^t$, where $\delta_1 = \frac{\epsilon^2}{256}$ is a small constant. In other words, Bob performs list-decoding on the left mega sub-code w.r.t. t, i.e. $\mathcal{C}_1 \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_k$, with a list-decoding radius equal to $r_{list} = t(\hat{p}_t \star q + \delta_1)$. Let the list of messages obtained in this phase be denoted by $\mathcal{L}$. We have,

$$\mathcal{L} = \{u \in \mathcal{U} : \exists\ (s_1, \cdots, s_k) \in \mathcal{S}^k\ s.t.\ d_H\left(\mathcal{C}_1(u, s_1) \circ \cdots \mathcal{C}_k(u, s_k), \mathbf{y}_1^t\right) \leq t(\hat{p}_t \star q + \delta_1)\}.$$

Unique decoding: In the unique decoding phase, Bob forms the set $\mathcal{A}$ of all possible right mega sub-codewords w.r.t. t (one for each possible sequence of secrets $s_{k+1}, s_{k+2}, \cdots, s_{1/\theta}$) for each message u in the list $\mathcal{L}$, i.e.,

$$\mathcal{A} = \{\mathcal{C}_{k+1}(u, s_{k+1}) \circ \mathcal{C}_{k+2}(u, s_2) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u, s_{\frac{1}{\theta}}) : u \in \mathcal{L}, (s_{k+1}, \cdots, s_{1/\theta}) \in \mathcal{S}^{\frac{1}{\theta}-k}\}.$$

He then considers Hamming balls of radius $r_{unique} = (n-t)\left(\frac{1-\theta}{4} + \frac{q(1+\theta)}{2}\right) = (n-t)\left(\frac{1}{4} \star q - \frac{\theta(1-2q)}{4}\right)$, each centered at a right mega sub-codeword from $\mathcal{A}$.

- If $\mathbf{y}_{t+1}^n$ lies within *exactly* one of the balls, the decoder outputs the message u' corresponding to its center, i.e., $\Gamma(\mathbf{y}) = u'$.

- If $\mathbf{y}_{t+1}^n$ lies in more than one ball, a decoding error is declared.
- If $\mathbf{y}_{t+1}^n$ lies outside all the balls, Bob picks the next chunk end in $\mathcal{T}$ and re-attempts decoding.

As we will show, depending on the adversary's attack strategy and the noise due to the BSC, there is a value of $t = t^*$ for which the decoding-attempt successfully recovers the transmitted message. However, Bob does know this value a priori. Bob begins by first identifying the smallest value of $t \geq n(1-4p')$ that coincides with a chunk end in $\mathcal{T}$, denoted $t_0 \in \mathcal{T}$, and performs a decoding attempt w.r.t t_0. Clearly, $t_0 = \min\{t : t \geq n(1-4p'), t \in \mathcal{T}\} = \left\lceil \frac{1-4p'}{\theta} \right\rceil n\theta$. If no message is returned, he re-attempts decoding with the next chunk end, $t = t_0 + n\theta$, and so on, each time picking a chunk end from the set $\mathcal{T}' = \{t_0, t_0 + n\theta, \cdots, n - n\theta\}$ until a message is returned. At any point in the decoding process, if $\mathbf{y}_{t+1}^n$ during unique decoding lies in more than one ball, a decoding error is declared and decoding terminates. If all decoding attempts fail to return a message having reached the end of the codeword, again a decoding error is declared.

Analysis: We begin our analysis with the following useful lemma.

Lemma 12. *Let $p, q \in [0, 1/2)$ and γ be a small positive constant such that $\gamma(1-2q) < 1/16$ and $p + \gamma < 1/2$. Then, we have the inequality $h_2\left((p+\gamma) \star q\right) < h_2(p \star q) + 2\sqrt{\gamma}$, where recall $x \star y = x(1-y) + y(1-x)$.*

Proof. Note $h_2\left((p+\gamma)\star q\right) \overset{(a)}{<} h_2(p\star q) + 2\gamma(1-2q)\log_2\left(\frac{1}{2\gamma - 4\gamma q}\right) \overset{(b)}{<} h_2(p\star q) + 2\sqrt{\gamma(1-2q)} \overset{(c)}{\leq} h_2(p \star q) + 2\sqrt{\gamma}$, where (a) follows from the inequality $h_2(a+b) < h_2(a) + 2b\log_2\left(\frac{1}{b}\right)$ (see for example [15, Lemma A.5] for a proof), (b) follows from the fact that $x \log_2\left(\frac{1}{x}\right) < \sqrt{x}$ when $x < \frac{1}{16}$ and (c) is true because $(1-2q) \in (0, 1]$. □

Reference trajectory $\hat{p}_t$: We now give an exact definition of $\hat{p}_t$, the reference trajectory for adversarial bit-flip attempts. It suffices to use the same $\hat{p}_t$ as defined in [15] where no BSC was present ($q = 0$) i.e. the decoder sets $\hat{p}_t$ independent of q.

Definition 1. (Definition of $\hat{p}_t$) Let $t \in \mathcal{T}$ be some chunk-end and recall $p' = p + \frac{\epsilon^2}{16}$. Define, $x_t = p' - \frac{(n-t)}{4n}$. For $t < n(1-4p')$, $\hat{p}_t = 0$. For $t \geq n(1-4p')$, $\hat{p}_t$ is defined to be

$$\hat{p}_t = \frac{x_t}{\alpha(p', x_t)} + \frac{\epsilon^2}{16\alpha^2(p', x_t)},$$

where $\alpha(p', x_t) = 1 - 4(p' - x_t) = \frac{t}{n}$.

In the following lemma, we prove that $\hat{p}_t$ satisfies two key technical conditions, the so-called *list decoding condition* given by (20), and the *energy bounding condition* given by (21).

Lemma 13. *(Modified from [15, Claim A.6]) For any* $t \in \mathcal{T}$ *such that* $t \geq n(1-4p')$*, the reference trajectory* $\hat{p}_t$ *satisfies*

$$t\left(1-h_2\left(\hat{p}_t \star q\right)\right)-\frac{n\epsilon}{2} \geq nR \tag{20}$$

and

$$pn - t\hat{p}_t \leq (n-t)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right). \tag{21}$$

Proof. Note that (21) follows directly from [15, Claim A.6] as it does not involve q. We only need to verify that (20) holds. Diving (20) by n and noting that $\alpha(p', x_t) = t/n$, we need to show that $\alpha(p',x_t)\left(1-h_2\left(\hat{p}_t \star q\right)\right)-\frac{\epsilon}{2} \geq R$. Substituting in the value of $\hat{p}_t$, we have

$$\begin{aligned}
&\alpha(p',x_t)\left(1-h_2\left(\left(\frac{x_t}{\alpha(p',x_t)}+\frac{\epsilon^2}{16\alpha^2(p',x_t)}\right)\star q\right)\right)-\frac{\epsilon}{2}\\
&\overset{(a)}{\geq} \alpha(p',x_t)\left(1-h_2\left(\frac{x_t}{\alpha(p',x_t)}\star q\right)-2\sqrt{\frac{\epsilon^2}{16\alpha^2(p',x_t)}}\right)-\frac{\epsilon}{2}\\
&= \alpha(p',x_t)\left(1-h_2\left(\frac{x_t}{\alpha(p',x_t)}\star q\right)\right)-\epsilon\\
&\geq \min_{x_t\in[0,p']}\alpha(p',x_t)\left(1-h_2\left(\frac{x_t}{\alpha(p',x_t)}\star q\right)\right)-\epsilon = C(p',q)-\epsilon = R,
\end{aligned}$$

proving the result, where inequality (a) follows from Lemma 12. □

Correct decoding point t^***:** From [15, Section A.3], for any trajectory p_t chosen by Calvin, Bob's reference trajectory $\hat{p}_t$ intersects p_t at some point before the second to last chunk end. In particular, there is a $t^* \in \mathcal{T}' = \{t_0, t_0+n\theta, \cdots, n-n\theta\}$ such that

$$\forall t \in \{t_0, t_0+n\theta, \cdots, t^*-n\theta\}, \qquad p_t > \hat{p}_t, \tag{22}$$

$$p_{t^*} \leq \hat{p}_{t^*}, \tag{23}$$

and

$$\forall t \in \{t_0, \cdots, t^*\}, \quad pn - tp_t \leq (n-t)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right). \tag{24}$$

As we will argue later, t^* defined above turns out to be the correct decoding point, where the two phase decoding attempt succeeds in finding the true message.

Key code properties: We now show that a code drawn at random from our ensemble satisfies with a positive probability two key properties.

List decoding property: This property will be used to prove that the size of the list obtained by Bob in a decoding attempt is at most a constant $O(1/\epsilon)$. We state it as the following lemma.

Lemma 14. *(Modified from [15, Claims A.15-A.16]) Suppose* $t \in \mathcal{T}' = \{t_0, t_0+n\theta, \cdots, n-n\theta\}$ *where* $t = kn\theta$ *satisfies* (20)*, i.e.* $t\left(1-h_2\left(\hat{p}_t \star q\right)\right) - \frac{n\epsilon}{2} \geq nR$*. Then, for sufficiently large* n*, with probability at least* $\left(1-\frac{1}{np}\right)$ *over the code design, the left mega sub-code* $\mathcal{C}_1 \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_k$ *is list decodable with radius* $r = t\left(\hat{p}_t \star q + \frac{\epsilon^2}{256}\right)$ *and list size* $L = O\left(\frac{1}{\epsilon}\right)$*.*

Proof. The proof follows the analysis in [15, Claims A.15-A.16]. The only additional step is to verify the bound

$$1-h_2\left(\hat{p}_t \star q + \frac{\epsilon^2}{256}\right) - \frac{nR}{t} - \frac{nS}{t\theta} \geq \frac{\epsilon}{4}.$$

Since $\theta = \frac{\epsilon^2(1-4p')}{4}$, $S = \frac{\theta^3}{8}$, from Lemma 12 and given that (20) is true, we have

$$1-h_2\left(\hat{p}_t \star q + \frac{\epsilon^2}{256}\right) - \frac{nR}{t} - \frac{nS}{t\theta} \geq \frac{n\epsilon}{2t} - \frac{n\theta^2}{8t} - 2\sqrt{\frac{\epsilon^2}{256}} \geq \frac{\epsilon}{4}$$

as desired. □

Distance property: For a decoding attempt at $t \in \mathcal{T}$, consider the list of messages $\mathcal{L}$ obtained by Bob in the list-decoding phase. Let u^* be the true message chosen by Alice for transmission and recall, $\mathcal{L}(u^*)$ is the set of all possible right mega sub-codewords w.r.t t for each message in $\mathcal{L} \setminus \{u^*\}$ i.e. $\mathcal{L}(u^*) = \{\mathcal{C}_{k+1}(u, s_{k+1}) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u, s_{\frac{1}{\theta}}) : u \in \mathcal{L}, u \neq u^*, (s_{k+1}, \cdots, s_{1/\theta}) \in \mathcal{S}^{\frac{1}{\theta}-k}\}$. Enumerate $\mathcal{L}(u^*)$ containing codewords of length $(n-t)$ as $\mathcal{L}(u^*) = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{|\mathcal{L}(u^*)|}\}$. The right mega-subcodeword for the true message is $\mathbf{x}_{t+1}^n(u^*, s_{right}) = \mathcal{C}_{k+1}(u^*, s_{k+1}) \circ \cdots \mathcal{C}_{\frac{1}{\theta}}(u^*, s_{\frac{1}{\theta}})$. We would like our code to satisfy the following distance condition

$$d_H\left(\mathbf{x}_{t+1}^n(u^*, s_{right}), \mathbf{w}_j\right) \geq (n-t)\left(\frac{1}{2} - \frac{\theta}{2}\right) \quad \forall \mathbf{w}_j \in \mathcal{L}(u^*). \tag{25}$$

Equation (25) is a key property that guarantees successful decoding. It ensures that the right mega sub-codeword for the transmitted message is sufficiently far in Hamming distance from the right mega sub words for any of the other messages in list $\mathcal{L}$. From [15, Claims A.20-A.23], (19) indeed occurs w.h.p., for almost all possible sequence of secrets s_{right}. We state this as the following lemma.

Lemma 15. *([15, Claims A.20-A.23]) Let* $C > 0$ *be an arbitrary constant. Then, there is a* n_0 *such that for* $n > n_0$*, with probability at least* $1-2^{-n}$*, a code drawn from the random ensemble satisfies the following property : for every chunk end* $t \in \mathcal{T}$*, for every message* u^**, and every list* $\mathcal{L}$ *of size at most* C/ϵ*, we have that* (25) *holds for at least a* $(1-2^{-nS/4})$ *portion of all possible secret sequences* s_{right}*.*

Success of decoding procedure: We are now ready to argue that the iterative decoding process succeeds in finding the true message with high probability. Fix a stochastic code $\mathcal{C} = \mathcal{C}_1 \circ \mathcal{C}_2 \circ \cdots \mathcal{C}_{1/\theta}$ for which both the list decoding property and the minimum distance property are satisfied, which we can do thanks to Lemmas 14 and 15. We will show that $t = t^*$ as defined by (22), (23) and (24) is in fact the correct decoding point i.e. at t^*, the list $\mathcal{L}$ obtained in the list decoding phase contains the true message which is then returned in the unique decoding phase.

Success of list decoding: When $t = t^*$, we have $\hat{p}_{t*} \geq p_{t^*}$. Thus, the number of adversarial bit-flip attempts injected onto $\mathbf{y}_1^{t^*}$, the left mega received word w.r.t. t^* is at most $t^*\hat{p}_{t^*}$. From Lemma 4 then, we have that $d_H(\mathbf{x}_1^{t^*}, \mathbf{y}_1^{t^*}) \leq t^* \left(\hat{p}_{t^*} \star q + \frac{\epsilon^2}{256}\right)$ with probability at least $1 - 2^{-\Omega(\epsilon^4 n)}$. Since the list-decoding radius is selected to be $r_{list} = t^* \left(\hat{p}_{t^*} \star q + \frac{\epsilon^2}{256}\right)$, the transmitted message is indeed in the list $\mathcal{L}$ with high probability as required.

Also note that when $t < t^*$, i.e., for $t \in \{t_0, t_0 + n\theta, \cdots, t^* - n\theta\}$, we have by the definition of t^* that $p_t > \hat{p}_t$. By a similar martingale argument as in Lemma 4 then, $\mathbf{y}_1^t$, the left mega received word w.r.t. t, lies w.h.p. outside the Hammming ball $\mathcal{B}(\mathbf{x}_1^t, r_{list})$. In other words, when $t < t^*$, the transmitted message u^* is w.h.p. not in the list $\mathcal{L}$ obtained by Bob.

Success of unique decoding: For $t_0 \leq t \leq t^*$, our code for almost all key sequences s_{right} satisfies

$$d_H\left(\mathbf{x}_{t+1}^n(u^*, s_{right}), \mathbf{w}_j\right) \geq (n-t)\left(\frac{1}{2} - \frac{\theta}{2}\right) \ \forall \mathbf{w}_j \in \mathcal{L}(u^*), \tag{26}$$

where recall that $\mathbf{w}_j$'s are the right-mega subcodewords corresponding to messages in $\mathcal{L}$ excluding u^*. Further, we have that Calvin has at most $(n-t)\left(\frac{1}{4} - \frac{\epsilon^2}{16}\right)$ bit-flip attempts left to inject onto $\mathbf{x}_{t+1}^n$. Recall also that Bob considers Hamming balls of radius $r_{unique} = (n - t)\left(\frac{1}{4} \star q - \frac{\theta(1-2q)}{4}\right)$ that are each centered at right-mega subcodewords in $\mathcal{L}$.

When $t_0 \leq t < t^*$, the true message $u^* \notin \mathcal{L}$ while at $t = t^*$ we have that $u^* \in \mathcal{L}$. At $t = t^*$, from Lemma 4, we have that for any adversarial strategy, $d(\mathbf{x}_{t^*+1}^n, \mathbf{y}_{t^*+1}^n) \leq (n - t^*)\left(\left(\frac{1}{4} - \frac{\epsilon^2}{16}\right) \star q + \gamma_1\right)$, with probability at least $1 - 2^{-\Omega(\gamma_1^2 n)}$. Choosing $\gamma_1 = \frac{p'\epsilon^2}{8}(1-2q)$, we have $d(\mathbf{x}_{t^*+1}^n, \mathbf{y}_{t^*+1}^n) \leq r_{unique}$. Thus, $\mathbf{y}_{t^*+1}^n$ is indeed w.h.p. inside the Hamming ball $\mathcal{B}(\mathbf{x}_{t^*+1}^n, r_{unique})$. Next, consider any $t_0 \leq t \leq t^*$ and $\mathbf{w}_j$ from the set $\mathcal{L}(u^*)$. We argue that as required, no matter what Calvin does, $\mathbf{y}_{t+1}^n$ is outside $\mathcal{B}(\mathbf{w}_j, r)$. Let $\mathcal{I}$ be the set of indices where $\mathbf{w}_j$ and $\mathbf{x}_{t+1}^n(u^*, s_{right})$ agree and $\mathcal{I}^c$ be the set of indices where they disagree. For a vector $\mathbf{v}$, let $(\mathbf{v})_{\mathcal{I}}$ denote $\mathbf{v}$ restricted to indices from $\mathcal{I}$. We have that

$$d_H(\mathbf{x}_{t+1}^n, \mathbf{y}_{t+1}^n) = d_H\left((\mathbf{x}_{t+1}^n)_{\mathcal{I}}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}}\right) + d_H\left((\mathbf{x}_{t+1}^n)_{\mathcal{I}^c}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}^c}\right) \tag{27}$$

and

$$d_H(\mathbf{w}_j, \mathbf{y}_{t+1}^n) = d_H\left((\mathbf{w}_j)_{\mathcal{I}}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}}\right) + d_H\left((\mathbf{w}_j)_{\mathcal{I}^c}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}^c}\right). \tag{28}$$

Now, Bob decodes $\mathbf{y}_{t+1}^n$ incorrectly to $\mathbf{w}_j$ when $d_H(\mathbf{x}_{t+1}^n, \mathbf{y}_{t+1}^n) > r_{unique}$ and $d_H(\mathbf{w}_j, \mathbf{y}_{t+1}^n) \leq r_{unique}$. Calvin's desire is then to inject his remaining bit-flip attempts in such a way that $\mathbf{y}_{t+1}^n$ is as far away as possible from $\mathbf{x}_{t+1}^n$, and at the same time, as close as possible to $\mathbf{w}_j$. Clearly, the best strategy is to only inject bit-flip attempts onto $(\mathbf{x}_{t+1}^n)_{\mathcal{I}^c}$. Then, since $(\mathbf{x}_{t+1}^n)_{\mathcal{I}}$ only suffers corruption due to the BSC(q), by the Chernoff bound we have

$$|\mathcal{I}|(q-\eta_1) \leq d_H\left((\mathbf{x}_{t+1}^n)_{\mathcal{I}}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}}\right) \leq |\mathcal{I}|(q+\eta_1) \tag{29}$$

with probability at least $(1-2^{-\Omega(\eta_1^2 n)})$. By Lemma 4 for $\mathcal{I}^c$, we also have

$$d_H\left((\mathbf{x}_{t+1}^n)_{\mathcal{I}^c}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}^c}\right) \leq |\mathcal{I}^c|\left(\left(\frac{(n-t^*)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right)}{|\mathcal{I}^c|}\right)\star q + \eta_2\right) \tag{30}$$

with probability at least $(1-2^{-\Omega(\eta_2^2 n)})$. By definition of $\mathcal{I}$ and $\mathcal{I}^c$, (29) and (30) then imply that

$$d_H\left((\mathbf{w}_j)_{\mathcal{I}}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}}\right) \geq |\mathcal{I}|(q-\eta_1) \tag{31}$$

and

$$d_H\left((\mathbf{w}_j)_{\mathcal{I}^c}, (\mathbf{y}_{t+1}^n)_{\mathcal{I}^c}\right) \geq |\mathcal{I}^c|\left(1-\left(\frac{(n-t)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right)}{|\mathcal{I}^c|}\right)\star q - \eta_2\right). \tag{32}$$

Consider the worst case when (26) holds with equality, i.e. $|\mathcal{I}^c|= (n-t)\left(\frac{1}{2}-\frac{\theta}{2}\right)$, and $|\mathcal{I}|= (n-t)\left(\frac{1}{2}+\frac{\theta}{2}\right)$. Since $(n-t)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right) < \frac{|\mathcal{I}^c|}{2}$, there is a constant $\delta_1 > 0$ that is only of ϵ, q and p such that $\frac{(n-t)\left(\frac{1}{4}-\frac{\epsilon^2}{16}\right)}{|\mathcal{I}^c|}\star q = \frac{\frac{1}{4}-\frac{\epsilon^2}{16}}{\frac{1}{2}-\frac{\theta}{2}}\star q = \left(\frac{1}{2}-\delta_1\right)$. We have then from (28) that $d_H(\mathbf{w}_j, \mathbf{y}_{t+1}^n) \geq |\mathcal{I}|(q-\eta_1)+|\mathcal{I}^c|\left(\frac{1}{2}+\delta_1-\eta_2\right)$. We have also from (29) and (30) that $d_H(\mathbf{x}_{t+1}^n, \mathbf{y}_{t+1}^n) \leq |\mathcal{I}|(q+\eta_1)+ |\mathcal{I}^c|\left(\frac{1}{2}-\delta_1+\eta_2\right)$. Now, choosing for instance $\eta_1 = \eta_2 = \delta_1/4$, it is easy to check $|\mathcal{I}|(q+\eta_1) + |\mathcal{I}^c|\left(\frac{1}{2}-\delta_1+\eta_2\right) < r_{unique} < |\mathcal{I}|(q-\eta_1) + |\mathcal{I}^c|\left(\frac{1}{2}+\delta_1-\eta_2\right)$ which implies that w.h.p., we will have $d_H(\mathbf{x}_{t+1}^n, \mathbf{y}_{t+1}^n) \leq r_{unique}$ and $d_H(\mathbf{w}_j, \mathbf{y}_{t+1}^n) > r_{unique}$. The argument holds for any $\mathbf{w}_j \in \mathcal{L}(u^*)$. Summarising, we have that w.h.p., no matter the strategy selected by Calvin,

- when $t_0 \leq t < t^*$, the transmitted message u^* is not in the list obtained by Bob, and $d(\mathbf{y}_{t+1}^n, \mathbf{w}_j) > r_{unique}$ for all $\mathbf{w}_j \in \mathcal{L}(u^*)$.
- when $t = t^*$, the transmitted message u^* is indeed in the list obtained by Bob. Further, we have $d(\mathbf{y}_{t+1}^n, \mathbf{x}_{t+1}^n) \leq r_{unique}$ and $d(\mathbf{y}_{t+1}^n, \mathbf{w}_j) > r_{unique}$ for all $\mathbf{w}_j \in \mathcal{L}(u^*)$.

Thus, the iterative decoding procedure used by Bob succeeds in finding the true message u^*.

VI. Conclusion

Motivated by security aspects for 5G networks and beyond, we considered the problem of communicating a message reliably through a BEC(q) or a BSC(q) with an adversary present who causally snoops in on both communicating parties and injects up to pn additional erasures or flips respectively. We gave a tight capacity characterization for each case. There are several interesting questions that remain open. Our achievability results prove only the existence of capacity-achieving stochastic codes. It is not even known whether stochastic codes are necessary to achieve capacity. In either case, it is desirable to find practical coding schemes with efficient encoding and decoding. One interesting research direction also is to characterize capacity when Calvin cannot snoop on Bob, i.e., feedback snooping is absent. Another is to characterize capacity when Alice has feedback and employs closed-loop encoding.

References

[1] V. Suresh, E. Ruzomberka, and D. J. Love, “Stochastic-adversarial channels : Online adversaries with feedback snooping,” in *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2021.

[2] ——, “Stochastic-adversarial channels : Online adversaries with feedback snooping,” 2021. [Online]. Available: https://arxiv.org/abs/2104.07194

[3] J. Cao, M. Ma, H. Li, R. Ma, Y. Sun, P. Yu, and L. Xiong, “A survey on security aspects for 3GPP 5G networks,” *IEEE Communications Surveys Tutorials*, vol. 22, no. 1, pp. 170–195, 2020.

[4] R. S. Chakraborty, S. Narasimhan, and S. Bhunia, “Hardware trojan: Threats and emerging solutions,” in *2009 IEEE International High Level Design Validation and Test Workshop*, 2009, pp. 166–171.

[5] Y. Jin and Y. Makris, “Hardware trojans in wireless cryptographic ICs,” *IEEE Design Test of Computers*, vol. 27, no. 1, pp. 26–35, 2010.

[6] K. S. Subramani, N. Helal, A. Antonopoulos, A. Nosratinia, and Y. Makris, “Amplitude-modulating analog/RF hardware trojans in wireless networks: Risks and remedies,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3497–3510, 2020.

[7] A. Lapidoth and P. Narayan, “Reliable communication under channel uncertainty,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, 1998.

[8] M. Langberg, “Oblivious communication channels and their capacity,” *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 424–429, 2008.

[9] V. Guruswami and A. Smith, “Codes for computationally simple channels: Explicit constructions with optimal rate,” in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, 2010, pp. 723–732.

[10] I. Csiszar and P. Narayan, “Arbitrarily varying channels with constrained inputs and states,” *IEEE Transactions on Information Theory*, vol. 34, no. 1, pp. 27–34, 1988.

[11] ——, “The capacity of the arbitrarily varying channel revisited: positivity, constraints,” *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, 1988.

[12] E. N. Gilbert, “A comparison of signalling alphabets,” *The Bell System Technical Journal*, vol. 31, no. 3, pp. 504–522, 1952.

[13] R. R. Varshamov, “Estimate of the number of signals in error correcting codes,” *Docklady Akad. Nauk, SSSR*, vol. 117, pp. 739–741, 1957.

[14] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, “New upper bounds on the rate of a code via the delsarte-macwilliams inequalities,” *IEEE Transactions on Information Theory*, vol. 23, no. 2, pp. 157–166, 1977.

[15] Z. Chen, S. Jaggi, and M. Langberg, “A characterization of the capacity of online (causal) binary channels,” in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 2015, pp. 287–296.

[16] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, “Upper bounds on the capacity of binary channels with causal adversaries,” *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3753–3763, 2013.

[17] R. Bassily and A. Smith, “Causal erasure channels,” in *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 2014, pp. 1844–1857.

[18] B. K. Dey, S. Jaggi, and M. Langberg, “Codes against online adversaries: Large alphabets,” *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3304–3316, 2013.

[19] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, “Coding against delayed adversaries,” in *2010 IEEE International Symposium on Information Theory*, 2010, pp. 285–289.

[20] ——, “A bit of delay is sufficient and stochastic encoding is necessary to overcome online adversarial erasures,” in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 880–884.

[21] B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, “The interplay of causality and myopia in adversarial channel models,” in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 1002–1006.

[22] B. K. Dey, S. Jaggi, and M. Langberg, “Sufficiently myopic adversaries are blind,” *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5718–5736, 2019.

[23] M. Langberg, “Private codes or succinct random codes that are (almost) perfect,” in *45th Annual IEEE Symposium on Foundations of Computer Science*, 2004, pp. 325–334.

[24] A. Smith, “Scrambling adversarial errors using few random bits, optimal information reconciliation, and better private codes,” in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA ’07. USA: Society for Industrial and Applied Mathematics, 2007, p. 395–404.

[25] S. Bhattacharya, A. J. Budkuley, and S. Jaggi, “Shared randomness in arbitrarily varying channels,” in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 627–631.

[26] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[27] N. Alon and J. H. Spencer, *The probabilistic method*. John Wiley & Sons, 2016.

[28] M. Langberg, S. Jaggi, and B. K. Dey, “Binary causal-adversary channels,” in *2009 IEEE International Symposium on Information Theory*. IEEE, 2009, pp. 2723–2727.

[29] A. Barg and G. Forney, “Random codes: minimum distances and error exponents,” *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2568–2573, 2002.

Appendix A
Proof of Lemma 9

From (17), we have that with probability at least $1-2^{-\Omega(\delta^2 n)}$, $\hat{\lambda}_{t^*} = \lambda_{t^*} - qt^* \in [\lambda^a_{t^*}(1-q+\delta) - \delta t^*, \lambda^a_{t^*}(1-q-\delta)+\delta t^*]$. We prove the lemma by showing that a small enough $\delta > 0$ can be set so that a t^* satisfying both conditions (15), (16) exists at the extremes $\hat{\lambda}_{t^*} = \lambda^a_{t^*}(1-q) - \delta(t^* - \lambda^a_{t^*})$ and $\hat{\lambda}_{t^*} = \lambda^a_{t^*}(1-q) + \delta(t^* - \lambda^a_{t^*})$. In the first case, we need to prove existence of t^* such that

$$n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 + \frac{\delta}{1-q}\right) + \theta t^* \leq t^*\left(1 + \frac{\delta}{1-q}\right) \tag{33}$$

and

$$np - \lambda^a_{t^*} \leq \frac{(n-t^*)(1-\theta)}{2} - \frac{\delta}{1-q}(t^* - \lambda^a_{t^*}). \tag{34}$$

First, choose a $t^* \leq n - n\theta$ in $\mathcal{T}$ such that $t^* \geq n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 + \frac{\delta}{1-q}\right) + \left(\theta - \frac{\delta}{1-q}\right)(n - n\theta)$. This ensures that (33) holds. Rearranging (34), we also require $t^* \leq \frac{n\left(1-\frac{2p}{1-\theta}\right)+\frac{2\lambda^a_{t^*}}{1-\theta}\left(1+\frac{\delta}{1-q}\right)}{1+\frac{2\delta}{(1-q)(1-\theta)}}$. Hence, to prove existence of t^* simultaneously satisfying both required conditions, it is sufficient to show that $\left(\frac{n\left(1-\frac{2p}{1-\theta}\right)+\frac{2\lambda^a_{t^*}}{1-\theta}\left(1+\frac{\delta}{1-q}\right)}{1+\frac{2\delta}{(1-q)(1-\theta)}}\right) - \left(n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 + \frac{\delta}{1-q}\right) + \left(\theta - \frac{\delta}{1-q}\right)(n - n\theta)\right) \geq n\theta$. Multiplying by $1 + \frac{2\delta}{(1-q)(1-\theta)}$ and simplifying, the coefficient of $\lambda^a_{t^*}$ in the above inequality becomes $\left(1 + \frac{\delta}{1-q}\right)\left(\frac{2}{1-\theta} - 1 - \frac{2\delta}{(1-q)(1-\theta)}\right)$ which is positive when $\delta < \frac{1}{2}(1+\theta)(1-q)$. For such a choice of δ, it is sufficient to show

$$p \leq \frac{1}{2}\left(\frac{1-\theta}{\theta}\right)\left(\epsilon - 2\theta + \theta^2 + \frac{\delta}{1-q}\left[1 - \theta - \frac{2}{1-\theta}\right]\right). \tag{35}$$

Since $\epsilon = 4\theta$, choosing $\delta < \min\left\{\frac{(\theta^2-\theta^3)(1-q)}{1+2\theta-\theta^2}, \frac{1}{2}(1+\theta)(1-q)\right\}$, we will have $\left(\frac{1-\theta}{\theta}\right)\left(\epsilon - 2\theta + \theta^2 + \frac{\delta}{1-q}[1 - \theta - \frac{2}{1-\theta}]\right) > 1$ so that (35) always holds for any $p \in [0, 1/2)$ and we are done. In the second case, we need to prove existence of t^* such that

$$n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 - \frac{\delta}{1-q}\right) + \theta t^* \leq t^*\left(1 - \frac{\delta}{1-q}\right) \tag{36}$$

and

$$np - \lambda^a_{t^*} \leq \frac{(n-t^*)(1-\theta)}{2} + \frac{\delta}{1-q}(t^* - \lambda^a_{t^*}). \tag{37}$$

Proceeding like earlier, choose a $t^* \leq n - n\theta$ in $\mathcal{T}$ such that $t^* \geq n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 - \frac{\delta}{1-q}\right) + \left(\theta + \frac{\delta}{1-q}\right)(n - n\theta)$, ensuring (36) holds. For (37) to hold, we need $t^* \leq \frac{n\left(1-\frac{2p}{1-\theta}\right)+\frac{2\lambda^a_{t^*}}{1-\theta}\left(1-\frac{\delta}{1-q}\right)}{1-\frac{2\delta}{(1-q)(1-\theta)}}$. Since the denominator $1 - \frac{2\delta}{(1-q)(1-\theta)} > 0$ for $\delta < \frac{1}{2}(1-\theta)(1-q)$, we will require that $n\left(1 - \frac{2p}{1-\theta}\right) + \frac{2\lambda^a_{t^*}}{1-\theta}\left(1 - \frac{\delta}{1-q}\right) - \left(n(1-2p-\epsilon) + \lambda^a_{t^*}\left(1 - \frac{\delta}{1-q}\right) + \left(\theta + \frac{\delta}{1-q}\right)(n - n\theta)\right) > n\theta$. Now, the coefficient of $\lambda^a_{t^*}$ in the above expression is $\frac{1+\theta}{1-\theta}\left(1 - \frac{\delta}{1-q}\right)$, which is always positive. Thus, we only need $p \leq \frac{1}{2}\left(\frac{1-\theta}{\theta}\right)\left(\epsilon - 2\theta + \theta^2 - \frac{\delta}{1-q}(1-\theta)\right)$. Proceeding exactly like before and choosing $\delta < (1-q)\frac{\theta^2}{1-\theta}$, this inequality always holds. Backtracking the proof steps, if we choose $\delta = \frac{1}{4}\frac{(1-q)\theta^2(1-\theta)}{1+2\theta-\theta^2}$, all of the required conditions are satisfied and the proof of this lemma is complete.

APPENDIX B
FORM OF $C^{flip}(p,q)$

Fix a $q \in [0,1/2)$. The optimization problem (5) in Theorem 2 is $\min_{0 \leq x \leq p} f(x)$ where $f(x) = (1-4p+4x)\left(1-h_2\left(\frac{x}{1-4p+4x} \star q\right)\right)$. When $p = 1/4$, $f(x) = 0$ at $x = 0$ and hence $C^{flip}(p,q) = 0$ when $p = 1/4$. Differentiating the objective function $f(x)$ we get $4 + (2q + 1)\log_2\left(\frac{x(1+2q)+q(1-4p)}{1-4p+4x}\right) + (3-2q)\log_2\left(\frac{1-4p+4x-x(1+2q)-q(1-4p)}{1-4p+4x}\right) = 0$. Solution x^* has the form $x^* = \frac{1-4p}{\alpha-3}$ where α satisfies $4 + (1+2q)\log_2\left(\frac{1-q(1-\alpha)}{1+\alpha}\right) + (3-2q)\log_2\left(\frac{\alpha+q(1-\alpha)}{1+\alpha}\right) = 0$. Since $0 \leq x \leq p$, we must have $\frac{1-4p}{\alpha-3} \leq p \implies p \geq \frac{1}{1+\alpha} = p_q$. Thus, for $p \in [p_q, 1/4]$, $x^* = \frac{(1-4p)p_q}{1-4p_q}$ where p_q satisfies

$$4 + (1+2q)\log_2\left(p_q \star q\right) + (3-2q)\log_2\left(1 - p_q \star q\right) = 0, \tag{38}$$

and the capacity expression becomes $C^{flip}(p,q) = \frac{1-4p}{1-4p_q}\left(1 - h_2(p_q \star q)\right)$. Thus, $C^{flip}(p,q)$, $p_q \leq p \leq 1/4$ is a straight line that intersects the p-axis at $p = 1/4$. For $p \in [0, p_q]$, the minimizer is $x^* = p$ and the capacity expression is $C^{flip}(p,q) = 1 - h_2(p \star q)$. Next we show that, $C^{flip}(p,q)$, $p_q \leq p \leq 1/4$ is in fact the tangent to the curve $1 - h_2(p \star q)$ at $p = p_q$. Consider the line $L(p)$ that is tangent to $1 - h_2(p \star q)$ and passes through $(1/4, 0)$. Its equation can be written as $L(p) = \gamma(1-4p)$ where γ is a constant. Suppose that $L(x)$ intersects $1 - h_2(p \star q)$ at $p = \tilde{p_q}$. To complete the proof, it suffices to show that $\tilde{p_q} = p_q$ i.e. $\tilde{p_q}$ satisfies (38). Since $L(p)$ is the tangent to $1 - h_2(p,q)$ at $p = \tilde{p_q}$, we have $\frac{d}{dp}L(p)\Big|_{p=\tilde{p_q}} = \frac{d}{dp}(1 - h_2(p \star q))\Big|_{p=\tilde{p_q}}$, which gives

$$-4\gamma = (1-2q)\log_2\left(\frac{\tilde{p_q} \star q}{1 - \tilde{p_q} \star q}\right). \tag{39}$$

We also have

$$L(\tilde{p_q}) = 1 - h_2(\tilde{p_q} \star q) = \gamma(1 - 4\tilde{p_q}). \tag{40}$$

Eliminating γ from (39) and (40), $\tilde{p_q}$ satisfies the equation $(1-2q)\log_2\left(\frac{\tilde{p_q} \star q}{1-\tilde{p_q} \star q}\right) = -4\left(\frac{1-h_2(\tilde{p_q} \star q)}{1-4\tilde{p_q}}\right)$. Rearranging the terms, this simplifies to $4 + (1+2q)\log_2\left(\tilde{p_q} \star q\right) + (3-2q)\log_2\left(1 - \tilde{p_q} \star q\right) = 0$ which is the same as (38). Hence, $p_q = \tilde{p_q}$ and we are done.