The Reliability Function of Arbitrary Channels With and Without Feedback†

Po-Ning Chen⋆ *and Fady Alajaji*‡

⋆ Industrial Technology Research Institute
Hsinchu, Taiwan 310, R.O.C

‡ Department of Mathematics and Statistics
Queen's University, Kingston, Ontario K7L 3N6, Canada

ABSTRACT

The reliability function of arbitrary single-user communication channels is analyzed. A general formula for the partitioning upper bound to the reliability function of channels without feedback is first established. This bound is indeed a consequence of the recently derived general expression for the Neyman-Pearson type-II error exponent subject to an exponential bound on the type-I error probability. A general lower bound to the non-feedback channel reliability function is also obtained using Feinstein's Lemma. Finally, the above bounds are extended to arbitrary channels with feedback.

1. INTRODUCTION

We investigate the reliability function $E(R)$ of *arbitrary* single-user channels (not necessarily memoryless, stationary, ergodic, information stable, etc.). A generalized partitioning upper bound to $R(E)$ (the dual function of $E(R)$) for channels without feedback is first established. We denote the upper bound by $\overline{R}(E)$. This bound is indeed a consequence of the recently derived general expression for the Neyman-Pearson hypothesis testing type-II error exponent subject to an exponential bound on the type-I error probability [1].

A general lower bound to the non-feedback channel reliability function is also obtained using Feinstein's Lemma. It is shown that in general, there exists a *gap* between the channel capacity and $R(0+)$, and a necessary and sufficient condition for eliminating this gap is derived.

Finally, the above bounds on the channel reliability function are extended to arbitrary single-user channels with output feedback.

†The work of Chen (poning@e0sun3.ccl.itri.org.tw) was supported in part by the Industrial Technology Research Institute under Grant 37H4100, sponsored by the Ministry of Economic Affairs, R.O.C. The work of Alajaji (fady@polya.mast.queensu.ca) was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant OGP0183645.

2. RELIABILITY FUNCTION OF CHANNELS WITHOUT FEEDBACK

Consider an arbitrary single-user channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. Let

$$\mathbf{X} = \{X^n = (X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)})\}_{n=1}^{\infty}$$

denote the channel input process in the form of a sequence of finite dimensional distributions, and let

$$\mathbf{Y} = \{Y^n = (Y_1^{(n)}, Y_2^{(n)}, \ldots, Y_n^{(n)})\}_{n=1}^{\infty}$$

be the corresponding output process induced by $\mathbf{X}$ via the channel n-dimensional transition distribution given by

$$\{P_{Y^n|X^n} : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}.$$

A channel code with blocklength n and rate R consists of an *encoder* $f(\cdot)$

$$f : \{1, 2, \ldots, 2^{nR}\} \to \mathcal{X}^n,$$

and a *decoder* $g(\cdot)$

$$g : \mathcal{Y}^n \to \{1, 2, \ldots, 2^{nR}\}.$$

The encoder represents the message $V \in \{1, 2, \ldots, 2^{nR}\}$ with the codeword $f(V) = X^n = [X_1, X_2, \ldots, X_n]$ which is then transmitted over the channel; at the receiver, the decoder observes the channel output $Y^n = [Y_1, Y_2, \ldots, Y_n]$, and chooses as its estimate of the message $\hat{V} = g(Y^n)$ (cf. Figure 1). A decoding error occurs if $\hat{V} \neq V$. Assuming that V is uniformly distributed over $\{1, 2, \ldots, 2^{nR}\}$, the (average) probability of decoding error is then given by

$$\begin{aligned} P_e(n, R) &= \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} Pr\{g(Y^n) \neq V | V = k\} \\ &= Pr\{g(Y^n) \neq V\}. \end{aligned}$$

We say that a rate R is *achievable (admissible)* if there exists a sequence of codes with blocklength n and rate R such that

$$\lim_{n\to\infty} P_e(n, R) = 0.$$

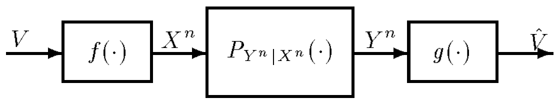

Figure 1. Block diagram of a channel without feedback.

The (non-feedback) capacity, C, of the channel is defined as the supremum of all achievable code rates.

In [2], Verdú and Han derived a general formula for the non-feedback capacity of arbitrary channels. It is shown to equal the supremum, over all input processes, of the input-output *inf-information rate* defined as the liminf in probability of the normalized information density:

$$C = \sup_{\mathbf{X}} \underline{I}(\mathbf{X};\mathbf{Y}),$$

where $\underline{I}(\mathbf{X};\mathbf{Y})$ is the *inf-information rate* between $\mathbf{X}$ and $\mathbf{Y}$ and is defined as the *liminf in probability* of the sequence of normalized information densities $\frac{1}{n}\, i_{X^nY^n}(X^n;Y^n)$, where

$$i_{X^nY^n}(a^n;b^n) = \log_2 \frac{P_{Y^n|X^n}(b^n|a^n)}{P_{Y^n}(b^n)}.$$

The *liminf in probability* of a sequence of random variables is defined as follows [2]: if A_n is a sequence of random variables, then its *liminf in probability* is the largest extended real number α such that for all $\xi > 0$,

$$\limsup_{n\to\infty} Pr[A_n \le \alpha - \xi] = 0.$$

Similarly, its *limsup in probability* is the smallest extended real numbers β such that for all $\xi > 0$,

$$\limsup_{n\to\infty} Pr[A_n \ge \beta + \xi] = 0.$$

Note that these two quantities are always defined; if they are equal, then the sequence of random variables converges in probability to a constant (which is α).

We now define the reliability function of channels without feedback, which we denote by $E(R)$. The channel reliability function is the *largest* exponent, among all possible (non-feedback) codes of rate R and codebook size $\exp\{nR\}$, of the probability of decoding error with respect to the codeword length n [3]. More specifically,

$$E(R) = \limsup_{n\to\infty} \sup_{X^n} -\frac{1}{n}\log P_e(n,R),$$

where the supremum is taken over all possible non-feedback codes. It follows by the definition of channel capacity (or the converse to the channel coding theorem) that $E(R) = 0$, for $R > C$. Furthermore, $E(R)$ is a non-increasing function of the rate R.

Theorem 1 ([1]) *Let $R(E)$ be the dual function of the channel reliability function $E(R)$ of an arbitrary channel (if it exists). Then for $E > 0$,*

$$\left(E + \sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D) < E\}\right) \vee 0 \ \ge\ R(E),$$

and

$$R(E) \ \ge\ \left(\sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D+E) < E\}\right) \vee 0,$$

where

$$L_{\mathbf{X}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log Pr\left\{\frac{1}{n}\, i_{X^nY^n}(X^n;Y^n) \le D\right\}.$$

The upper bound to $R(E)$ in Theorem 1 is the generalized version of the partitioning upper bound to the channel reliability function [3]. It is derived as a direct consequence of Theorem 3 in [1] which presents a general expression of the Neyman Pearson type-II error exponent subject to an exponential bound to the type-I error probability. This is achieved by transforming the channel reliability problem into a hypothesis testing problem and applying the result of Theorem 3 [1]. The lower bound to $R(E)$ in Theorem 1 is obtained using Feinstein's Lemma (Theorem 1 in [2]) which establishes the existence of a block code with a guaranteed decoding error probability as a function of its size. This bound constitutes the generalized version of the random coding lower bound to $E(R)$ presented in [3].

Lemma 1 ([1]) *Define*

$$\overline{R}(E) \triangleq E + \sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D) < E\}.$$

Then

$$\overline{R}(0^+) \ = \ \lim_{E\downarrow 0} \overline{R}(E) \le C = \sup_{\mathbf{X}} \underline{I}(\mathbf{X};\mathbf{Y}). \tag{2.1}$$

Furthermore, $R(0^+) = C$ if and only if

$$\forall\ \delta > 0, \qquad \sup_{\mathbf{X}} L_{\mathbf{X}}(C-\delta) > 0. \tag{2.2}$$

Observations:

- Note that for discrete memoryless channels (DMC), the partitioning upper bound to $R(E)$ is tight at rates close to C.
- For DMC's, $R(0^+) = C$. However, this is not true in general (cf. (2.1)). That is, there exits a *gap* between $R(0^+)$ and C for arbitrary channels.
- It is important to point out that condition (2.2) is *not* equivalent to information stability[1]. This is shown by the following counterexample.

[1]Information stable channels have the property that the input that maximizes mutual information and its corresponding output behave ergodically [2, 5]. The class of information stable channels is the most general class of channels for which the expression

$$C = \lim_{n\to\infty} \frac{1}{n} \sup_{(X_1,\ldots,X_n)} I\left((X_1,\ldots,X_n);(Y_1,\ldots,Y_n)\right)$$

represents the operational capacity.

Example: *The Polya Contagion Channel.* Consider the discrete binary additive channel of [6] described by:

$$Y_i = X_i \oplus Z_i, \qquad i = 1, 2, \cdots,$$

where the $\oplus$ represents the modulo-2 addition operation, random variables X_i, Z_i, and Y_i are, respectively, the i'th input, noise, and output of the channel. The input and noise sequences are assumed to be independent from each other. The noise sequence $\{Z_i\}$ is generated according to Polya's contagion urn scheme [6]. The Polya additive noise $\{Z_i\}$ forms an exchangeable (hence stationary) non-ergodic process [6]; resulting in an *information unstable* channel. It is shown in [6] that the channel capacity of this channel is $C = 0$. However, condition (2.2) holds for this channel. This is demonstrated in the following proposition.

Proposition 1 *Consider the Polya channel described above. Then,*

$$\forall\ \delta > 0, \qquad \sup_{\mathbf{X}} L_{\mathbf{X}}(-\delta) > 0,$$

which is equivalent to the fact that $R(0^+) = C = 0$.

Proof: Take an input process $\tilde{\mathbf{X}}$ to be Bernoulli(1/2) (i.e., IID and uniform):

$$Pr\{\tilde{X}^n = x^n\} = \left(\frac{1}{2}\right)^n, \quad \text{for all } x^n \in \{0,1\}^n.$$

Now recall that the channel large-deviation spectrum $L_{\tilde{\mathbf{X}}}(D)$ for input $\tilde{\mathbf{X}}$ is defined by:

$$L_{\tilde{\mathbf{X}}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log Pr\left\{\frac{1}{n}\, i_{\tilde{X}^n Y^n}(\tilde{X}^n; Y^n) \le D\right\}.$$

We can write:

$$\frac{1}{n} i_{\tilde{X}^n Y^n}(\tilde{X}^n; Y^n) = 1 + \frac{1}{n} log_2 P(Z^n).$$

So

$$\begin{aligned} Pr\left\{\frac{1}{n} i_{\tilde{X}^n Y^n}(\tilde{X}^n; Y^n) \le D\right\} &= Pr\left\{P(Z^n) \le 2^{n(D-1)}\right\} \\ &\le 2^{n(D-1)}\, 2^n \\ &= 2^{nD}. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{1}{n}\log Pr\left\{\frac{1}{n}\, i_{\tilde{X}^n Y^n}(\tilde{X}^n; Y^n) \le D\right\} &\le \frac{1}{n}\log_2 2^{nD} \\ &= D, \end{aligned}$$

and $L_{\tilde{\mathbf{X}}}(D) \ge -D$. Therefore, for all $\delta > 0$,

$$\sup_{\mathbf{X}} L_{\mathbf{X}}(C - \delta) = \sup_{\mathbf{X}} L_{\mathbf{X}}(-\delta) \ge L_{\tilde{\mathbf{X}}}(-\delta) \ge \delta > 0.$$

Condition (2.2) is hence satisfied and $R(0^+) = C = 0$. $\square$

Finally, we observe that the generalized partitioning upper bound to the channel reliability function in Theorem 1 can be tightened by replacing the *sup* operation by the *inf* operation in the upper bound expression. This yields the following result.

Corollary 1 *[Tighter upper bound] Let $R(E)$ be the dual function of the channel reliability function $E(R)$ of an arbitrary channel (if it exists). Then for $E > 0$,*

$$\left(E + \inf_{\mathbf{X}}\ \inf\{D : L_{\mathbf{X}}(D) < E\}\right) \vee 0 \ \ge\ R(E),$$

and

$$R(E)\ \ge\ \left(\sup_{\mathbf{X}}\ \inf\{D : L_{\mathbf{X}}(D+E) < E\}\right) \vee 0,$$

where

$$L_{\mathbf{X}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log Pr\left\{\frac{1}{n}\, i_{X^n Y^n}(X^n; Y^n) \le D\right\}.$$

3. RELIABILITY FUNCTION OF CHANNELS WITH FEEDBACK

We next consider arbitrary single-user channels with output feedback as depicted in Figure 2.

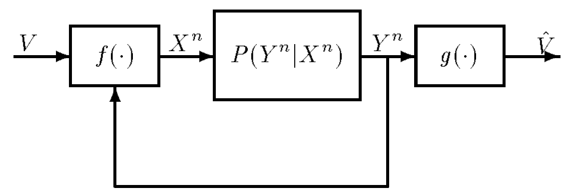

Figure 2. Block diagram of a channel with feedback.

As seen in the above diagram, there exists a "return" (or feedback) channel from the receiver to the transmitter; we assume this return channel is instantaneous, noiseless, and has large capacity. The receiver employs the feedback channel to inform the transmitter what letters were actually received; these letters are received at the transmitter before the next letter is sent, and therefore can be used in choosing the next transmitted letter.

A feedback code with blocklength n and rate R consists of sequence of encoders

$$f_i : \{1, 2, \ldots, 2^{nR}\} \times \mathcal{Y}^{i-1} \to \mathcal{X}$$

for $i = 1, 2, \ldots, n$, along with a decoding function

$$g:\ \mathcal{Y}^n\ \to \{1, 2, \ldots, 2^{nR}\},$$

where $\mathcal{X}$ and $\mathcal{Y}$ are the input and output alphabets, respectively. The interpretation is simple: If the user wishes to convey message $V \in \{1, 2, \ldots, 2^{nR}\}$ then the first code symbol transmitted is $X_1 = f_1(V)$; the second code symbol transmitted is $X_2 = f_2(V, Y_1)$, where Y_1 is the channel's output due to X_1. The third code symbol transmitted is $X_3 = f_3(V, Y_1, Y_2)$, where Y_2 is the channel's output due to X_2. This procedure is continued until the encoder transmits $X_n = f_n(V, Y_1, Y_2, \ldots, Y_{n-1})$. At this point the decoder estimates the message to be $g(Y^n)$, where $Y^n = [Y_1, Y_2, \ldots, Y_n]$.

Again, we assume that V is uniformly distributed over $\{1, 2, \ldots, 2^{nR}\}$, and we define the probability of error and achievability as in the previous section. The capacity of the channel with feedback, C_{FB}, is defined to be the supremum of all achievable feedback code rates.

In [4], the general capacity formula of arbitrary channels with feedback is derived. It is given by

$$C_{FB} = \sup_{\mathbf{X}} \underline{I}(\mathbf{V}; \mathbf{Y}),$$

where the supremum is taken over all possible feedback encoding schemes[2].

The channel reliability function with feedback, denoted by $E_{FB}(R)$, is the largest exponent, among all possible feedback codes of rate R and codebook size $\exp\{nR\}$, of the probability of decoding error with respect to the codeword length n. In other words,

$$E_{FB}(R) = \limsup_{n\to\infty} \sup_{X^n} -\frac{1}{n}\log P_e(n, R),$$

where the supremum is taken over all possible feedback codes.

We herein extend the results of the previous section for the case of channels with feedback.

Theorem 2 *Let $R_{FB}(E)$ be the dual function of the reliability function $E_{FB}(R)$ of an arbitrary channel with feedback (if it exists). Then for $E > 0$,*

$$\left(E + \sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D) < E\}\right) \vee 0 \geq R_{FB}(E),$$

and

$$R_{FB}(E) \geq \left(\sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D + E) < E\}\right) \vee 0,$$

where

$$L_{\mathbf{X}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log Pr\left\{\frac{1}{n} i_{VY^n}(V; Y^n) \leq D\right\},$$

and supremum is taken over all feedback codes.

Lemma 2 *Define*

$$\overline{R}_{FB}(E) \triangleq E + \sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D) < E\}.$$

Then

$$\overline{R}_{FB}(0^+) = \lim_{E\downarrow 0} \overline{R}_{FB}(E) \leq C_{FB} = \sup_{\mathbf{X}} \underline{I}(\mathbf{V}; \mathbf{Y}).$$

Furthermore, $R_{FB}(0^+) = C_{FB}$ if and only if

$$\forall\, \delta > 0, \qquad \sup_{\mathbf{X}} L_{\mathbf{X}}(C_{FB} - \delta) > 0. \tag{3.3}$$

[2]
$$\sup_{X^n} \underline{I}(V; Y^n) = \sup_{X^n = (f_1(V), f_2(V, Y_1), \ldots, f_n(V, Y^{n-1}))} \underline{I}(V; Y^n) = \sup_{(f_1, f_2, \ldots, f_n)} \underline{I}(V; Y^n).$$

Corollary 2 *[Tighter upper bound] Let $R_{FB}(E)$ be the dual function of the reliability function $E_{FB}(R)$ of an arbitrary channel with feedback (if it exists). Then for $E > 0$,*

$$\left(E + \inf_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D) < E\}\right) \vee 0 \geq R_{FB}(E),$$

and

$$R_{FB}(E) \geq \left(\sup_{\mathbf{X}} \inf\{D : L_{\mathbf{X}}(D + E) < E\}\right) \vee 0,$$

where

$$L_{\mathbf{X}}(D) = \liminf_{n\to\infty} -\frac{1}{n}\log Pr\left\{\frac{1}{n} i_{VY^n}(V; Y^n) \leq D\right\},$$

and supremum is taken over all feedback codes.

4. FUTURE WORK

Current efforts focus on the comparison of the general bounds to the channel reliability function derived in this paper with previously known bounds for the case of simple channels such as DMC's [3]. Another issue worth exploring is the effect of feedback on the channel reliability function.

Future work may address the analysis of the properties of $\overline{R}(E)$ and condition (2.2), and the determination of the class of channels for which $R(0^+) < C$.

REFERENCES

[1] P.-N. Chen, "General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds, " *IEEE Trans. Info. Theory*, vol. 42, pp. 316-323, January 1996.

[2] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Info. Theory*, vol. 40, pp. 1147-1157, July 1994.

[3] R. E. Blahut, *Principles and Practice of Information Theory*, Addison Wesley, Massachusetts, 1988.

[4] P.-N. Chen and F. Alajaji, "Strong converse, feedback channel capacity and hypothesis testing," *Journal of the Chinese Institute of Engineers*, vol. 18, pp. 777-785, November 1995; also in *Proceedings of CISS*, John Hopkins Univ., MD, USA, March 1995.

[5] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964.

[6] F. Alajaji and T. Fuja, "A communication channel modeled on contagion," *IEEE Trans. Info. Theory*, vol. 40, pp. 2035-2041, November 1994.