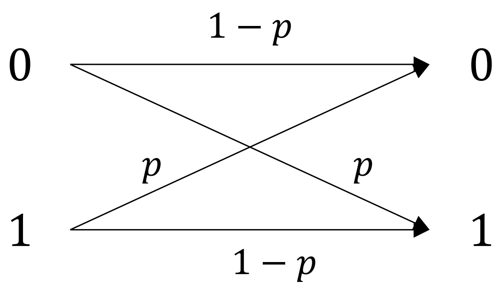
A Practical Coding Scheme for the BSC with Feedback

Ke Wu* and Aaron B. Wagner†
*Computer Science Department, Carnegie Mellon University
† School of Electrical and Computer Engineering, Cornell University.

Binary Symmetric Channel (BSC)

Encoder
$\mathcal{X} = \{0,1\}$



Decoder
$\mathcal{Y} = \{0,1\}$

$m \in \mathcal{M}$

Encoding mapping f ↓

$x^N = x_1, \dots, x_N \in \{0,1\}^N$

$y^N = y_1, \dots, y_N \in \{0,1\}^N$

↓ Decoding mapping g

m'

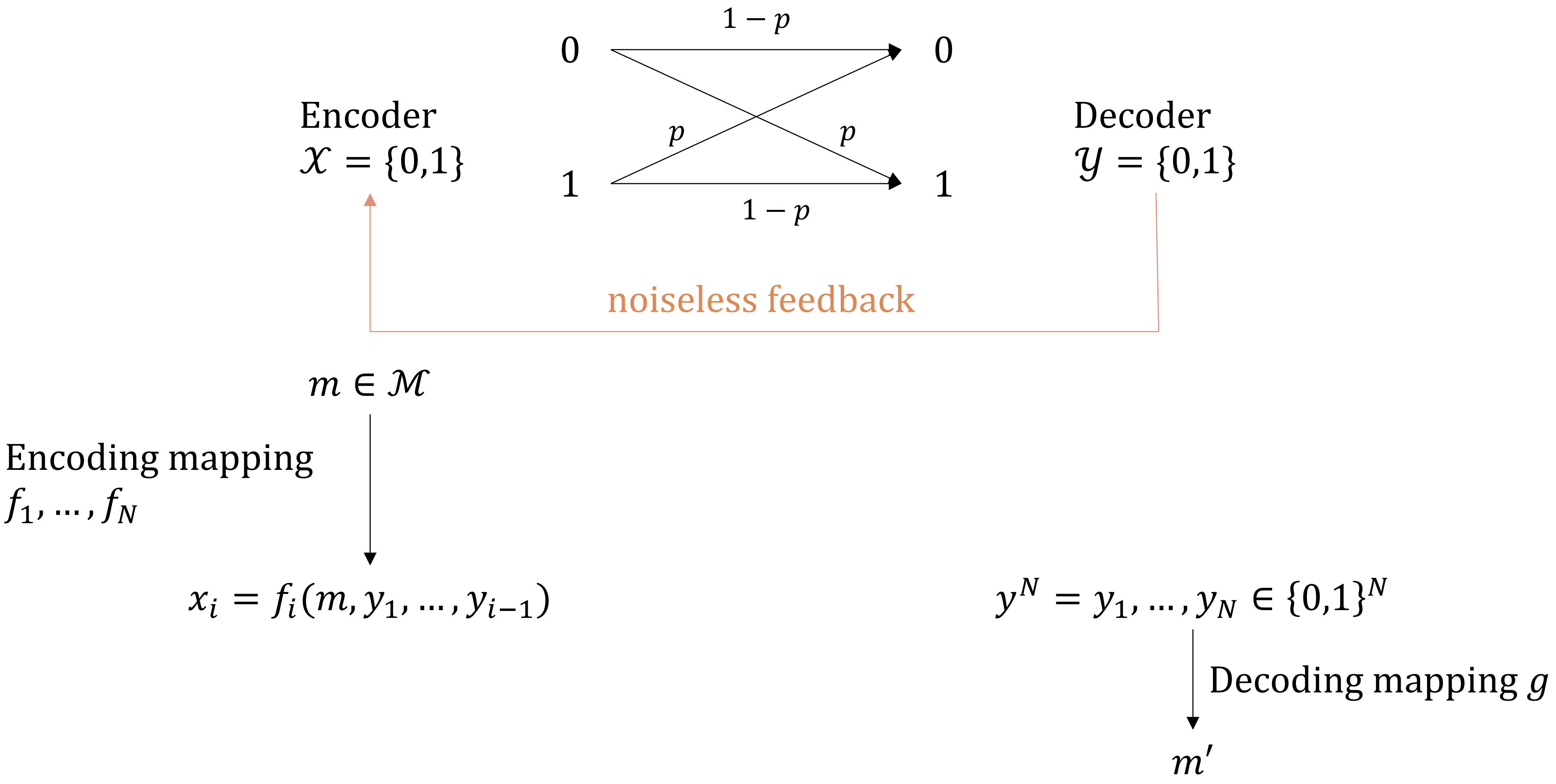
(f, g): code $\mathcal{C}_{N,R}$

N: block length

R: $\log \frac{|\mathcal{M}|}{N}$

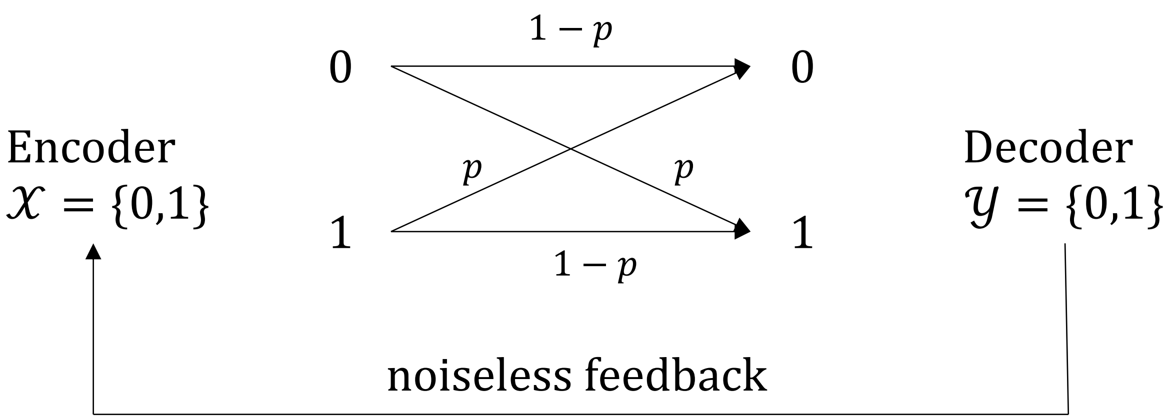
$P_e = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \Pr[m' \neq m]$

BSC with Noiseless Feedback



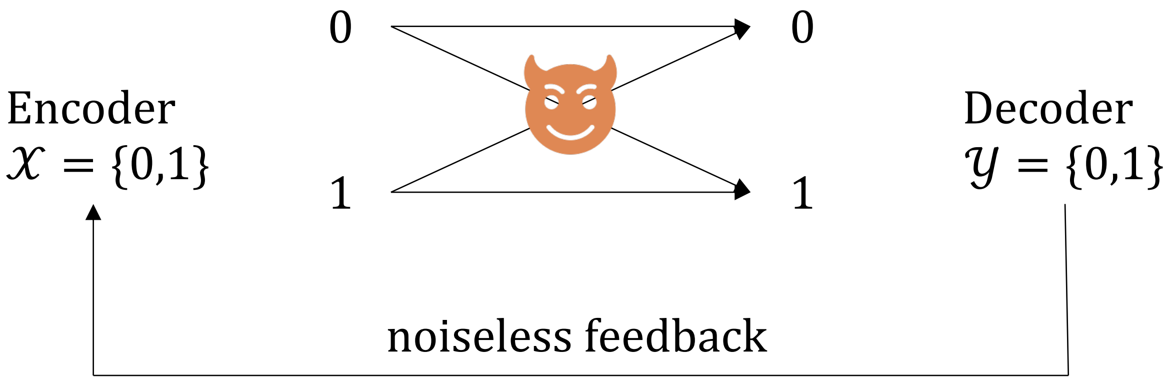
Feedback does not improve the channel capacity.
But it simplifies the coding.

BSC with Feedback: Stochastic vs. Adversarial



$BSC^{fb}(p)$

A sequence of codes $\{\mathcal{C}_{N,R}\}_N$ admissible for $BSC^{fb}(p)$ if $P_e \to 0$ as $N \to \infty$



$BSC_{adv}^{fb}(f)$

A code $\mathcal{C}$ admissible for $BSC_{adv}^{fb}(f)$ it can correct any f fraction of errors.

Existing Results

- Berlekamp [1] and Zigangirov [2] determined the capacity for adversarial BSC with feedback.
- Schalkwijk [3] proposed a simple scheme for adversarial BSC with feedback.
- Ahlswede *et al .* [4] then proposed a coding method called the rubber coding method for adversarial channel with feedback.

But none of these schemes are amenable to direct, efficient implementation.

- Horstein's scheme [5] and Li and El Gamal's scheme [6] can be implemented and can be applied to more general channels, but our scheme provides stronger optimality guarantees.

Our Results

- We propose a practically-implementable end-to-end scheme for BSC with feedback based on the rubber coding method;
- Our scheme is optimal with respect to the error-exponent, second order coding rate, and the moderate deviations performance for certain rates, and it is nearly optimal with respect to the third order coding rate and the pre-factor of the error exponent.
- Our analysis strictly enlarge the set of rates for which the sphere-packing bound is known to be achievable for the BSC with feedback.

Rubber Coding Method

Skeleton sequence: binary sequence with no ℓ consecutive zeros.

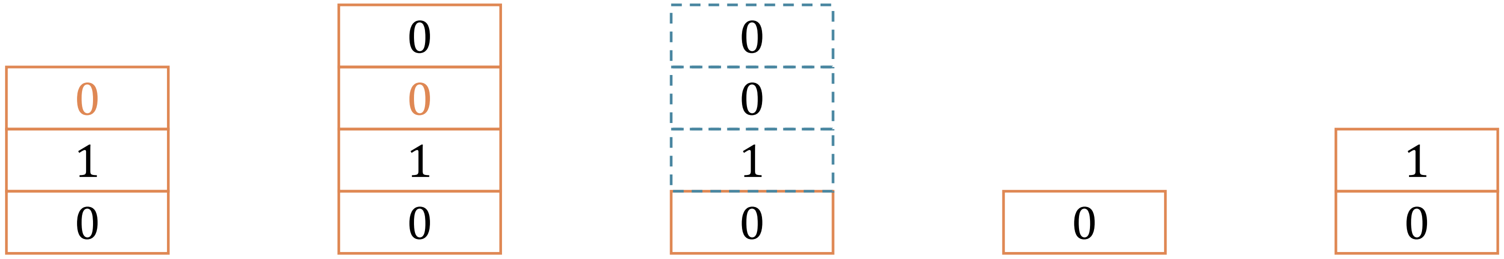
$\mathcal{A}_\ell^{N'}$: the set of all binary sequences of length N' with no ℓ consecutive zeros.

To send a skeleton sequence $x^{N'}$:

- Decoding: Whenever the decoder receives a bit, it checks if there are consecutive ℓ zeros. If yes, it removes the consecutive ℓ zeros as well as the bit before these consecutive ℓ zeros. Finally, it truncates the output to N' bits.
- Encoding: If the decoder's current version is a prefix of $x^{N'}$, send the next bit. If not, send 0.

Rubber Coding Method

Encoder wants to send a skeleton sequence $x = 011010 \in \mathcal{A}_2^6$



Decoder's stack

Need $\ell + 1$ bits to correct one error.

A rubber code with block length N is admissible for $BSC_{adv}^{fb}(f)$ if

$$N' + (\ell + 1) f N \leq N$$

Skeleton Sequences and Markov Chains

How do we map message $m \in \mathcal{M}$ to a skeleton sequence efficiently?

The stochastic process that is uniformly distributed over $\mathcal{A}_\ell^{N'}$ is an $(\ell - 1)$-th order Markov chain.

$$A_\ell(N) = \left|\mathcal{A}_\ell^N\right|$$

Initial distribution of the Markov chain:

$$\Pr[X_1 = 1] = \frac{A_2(N-1)}{A_2(N)}, \Pr[X_1 = 0] = 1 - \frac{A_2(N-1)}{A_2(N)}$$

Transition matrix of the Markov chain

$$\begin{array}{cc} & \overbrace{\begin{array}{cc} 0 & \quad\quad\quad 1 \end{array}}^{X_{i+1}} \\ X_i \left\{ \begin{array}{c} 0 \\ 1 \end{array} \right. & \begin{pmatrix} 0 & 1 \\ 1 - \dfrac{A_2(N-i-1)}{A_2(N-i)} & \dfrac{A_2(N-i-1)}{A_2(N-i)} \end{pmatrix} \end{array}$$

Skeleton Sequences and Markov Chains

$A_2(N) =$
- number of sequences begin with 1: $A_2(N-1)$.
- number of sequences begin with 0
  - number of sequences begin with 00: 0.
  - number of sequences begin with 01: $A_2(N-2)$.

$$A_2(N) = A_2(N-1) + A_2(N-2).$$

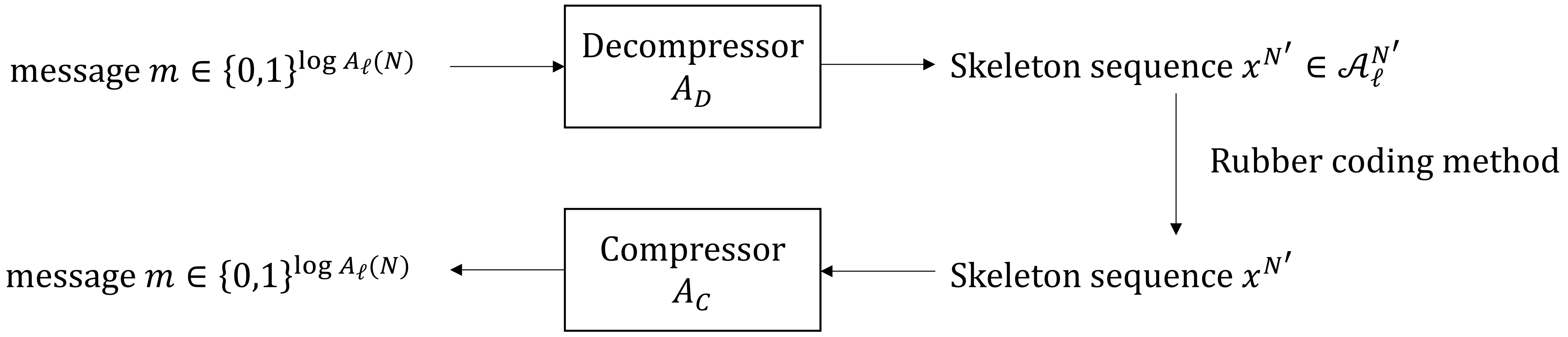
Initial condition: $A_2(1) = 2, A_2(2) = 3$.

The initial distribution and the transition matrix of the Markov chain that is uniform over the set of skeleton sequences can be efficiently computed.

Arithmetic Coding

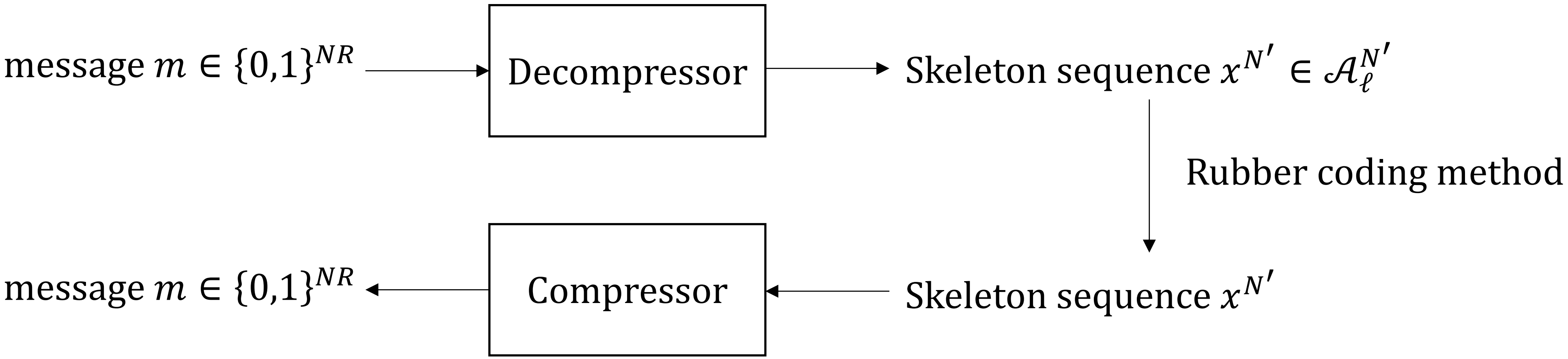
How do we map message $m \in \mathcal{M}$ to a skeleton sequence efficiently?

Shannon-Fano-Elias Code compresses a source Markov chain with known distribution to near optimal length.



Our Final Construction

$\mathcal{C}_{\ell,N,R}$:



Error probability: probability that the fraction of errors is larger than $\frac{1}{\ell+1}\left(1-\frac{N'}{N}\right)$

Streaming: the encoder does not need the entire message to begin encoding
Universal: the encoder does not need knowledge of the cross-over probability p

Optimality: Capacity Achieving

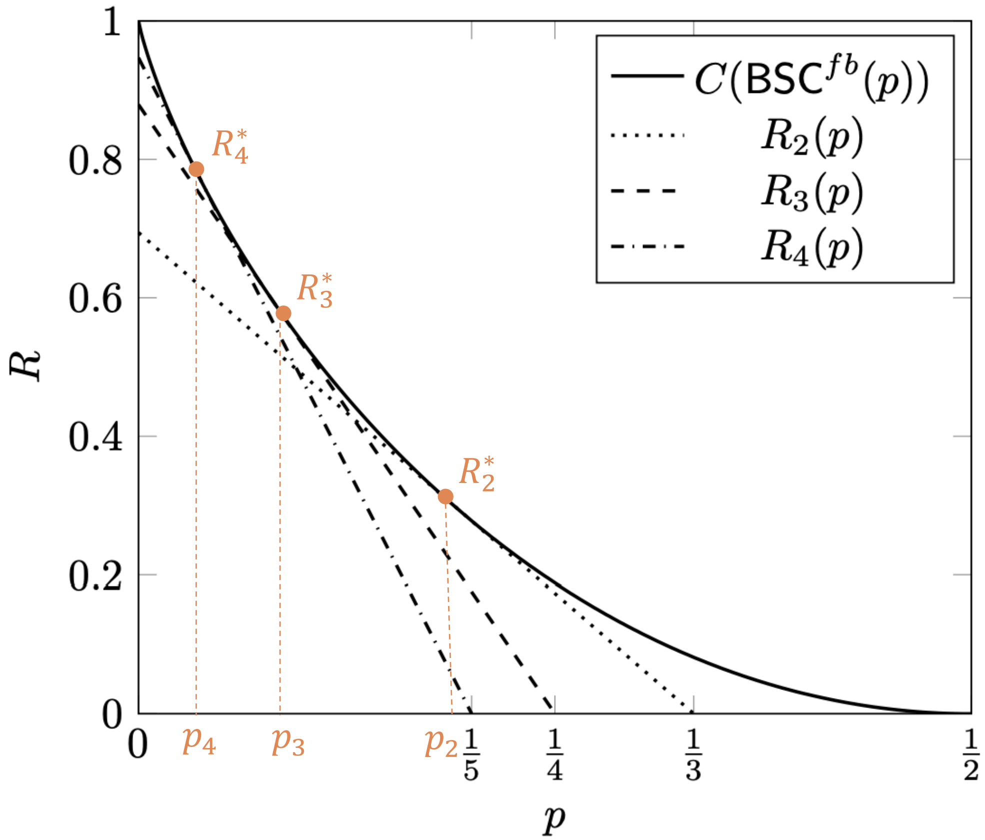


Fig. 3: $R_\ell(p)$ for different ℓ.

Optimality: Second-order Rate

Theorem *For any fixed* ℓ, $BSC^{fb}(p_\ell)$ *with cross-over probability* p_ℓ. $\forall \epsilon$, *let* $R(N,\epsilon)$ *denote the largest possible rate such that* $\mathcal{C}_{\ell,N,R(N,\epsilon)}$ *has error probability at most* ϵ, *and let* C *denote the capacity of* $BSC(p_\ell)$. *Then for large* N,

$$R(N,\epsilon) \geq C - \frac{1}{\sqrt{N}}\sqrt{p(1-p)\log^2\frac{1-p}{p}}\,\Phi^{-1}(1-\epsilon) - O\left(\frac{1}{N}\right).$$

(***Second-order coding rate***) *Given a block length* N *and an* ϵ *such that* $0 < \epsilon < 1$, *the largest possible rate of a code for the* $BSC^{f}(p)$ *with error probability less than or equal to* ϵ *is*

$$C - \frac{1}{\sqrt{N}}\sqrt{p(1-p)\log^2\frac{1-p}{p}}\,\Phi^{-1}(1-\epsilon) + \frac{\log N}{2N} + o(1).$$

Optimality: Moderate Deviations

Theorem *Fix any* $\ell \geq 2$. *Let* C *be the capacity of the* $BSC^f(p_\ell)$. *For any sequence of real numbers* ϵ_N *s.t.* $\epsilon_N \to 0$ *as* $N \to \infty$ *and* $\epsilon_N\sqrt{N} \to \infty$ *as* $N \to \infty$, *consider the sequence of codes* $\{\mathcal{C}_{\ell,N,R(N)}\}_N$ *such that* $R_N = C - \epsilon_N$. *Let* $P_e(\mathcal{C}_{\ell,N,R(N)})$ *denote the average error probability of* $\mathcal{C}_{\ell,N,R(N)}$ *over the* $BSC^f(p_\ell)$. *Then*

$$\lim_{N\to\infty} \frac{1}{N\epsilon_N^2} \log P_e\left(\mathcal{C}_{\ell,N,R(N)}\right) = -\frac{1}{2p(1-p)\log^2\frac{1-p}{p}},$$

(**Moderate deviations**) *For any sequence of real numbers* ϵ_N *s.t.* $\epsilon_N \to 0$ *as* $N \to \infty$ *and* $\epsilon_N\sqrt{N} \to \infty$ *as* $N \to \infty$, *for any sequence of codes* $\{\mathcal{C}_{N,R(N)}\}_N$ *such that* $R_N \geq C - \epsilon_N$, *we have*

$$\liminf_{N\to\infty} \frac{1}{N\epsilon_N^2} \log P_e\left(\mathcal{C}_{\ell,N,R(N)}\right) \geq -\frac{1}{2p(1-p)\log^2\frac{1-p}{p}}$$

Optimality: Error Exponent

Theorem For any fixed $\ell \geq\ 2$, consider the sequence of codes $\left\{\mathcal{C}_{\ell,N,R_\ell^*}\right\}_N$ at the tangent rate R_ℓ^* That is, $R_\ell^* \ = R_\ell(p_\ell) = 1 - h(p_\ell)$. Then for the $BSC^f(p)$with $p < p_\ell$, $\left\{\mathcal{C}_{\ell,N,R_\ell^*}\right\}_N$ at rate R_ℓ^*, the error probability

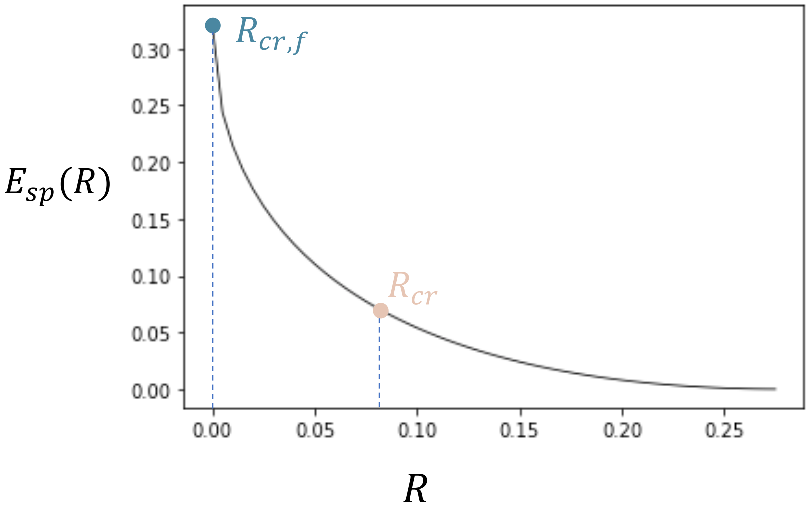
$$P_e\left(\mathcal{C}_{\ell,N,R_\ell^*}\right) \leq O\left(\frac{1}{\sqrt{N}}\right) e^{-N\cdot E_{sp}(R)}.$$

(**Sphere-packing bound with pre-factor**) *Let* $\mathcal{C}_{N,R}$ *be a sequence of codes for the* $BSC^f(p)$*, each with rate less* R_ℓ^* *than the capacity. Let* $E'_{sp}(R)$ *be the slope of* $E_{sp}(R)$ *at* R*. Then the error probability* $P_e\left(\mathcal{C}_{N,R}\right)$ *satisfies*

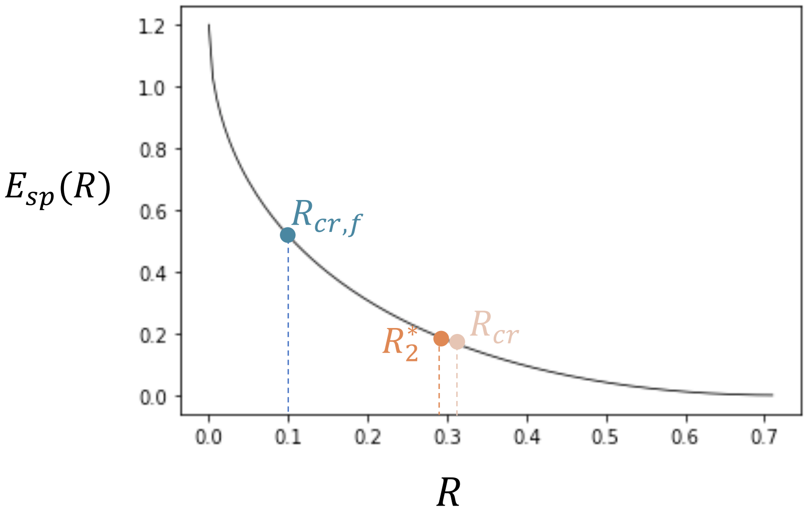
$$P_e\left(\mathcal{C}_{N,R}\right) \geq O\left(\frac{1}{N^{\frac{1}{2}\left(1+E'_{sp}(R)\right)}}\right) e^{-N\cdot E_{sp}(R)}.$$

Our “pre-factor” is $O\left(\frac{1}{\sqrt{N}}\right)$, while the optimal order of $O\left(\frac{1}{N^{\frac{1}{2}\left(1+E'_{sp}(R)\right)}}\right)$. For the binary erasure channel (BEC), both with and without feedback, the optimal pre-factor is $O\left(\frac{1}{\sqrt{N}}\right)$.

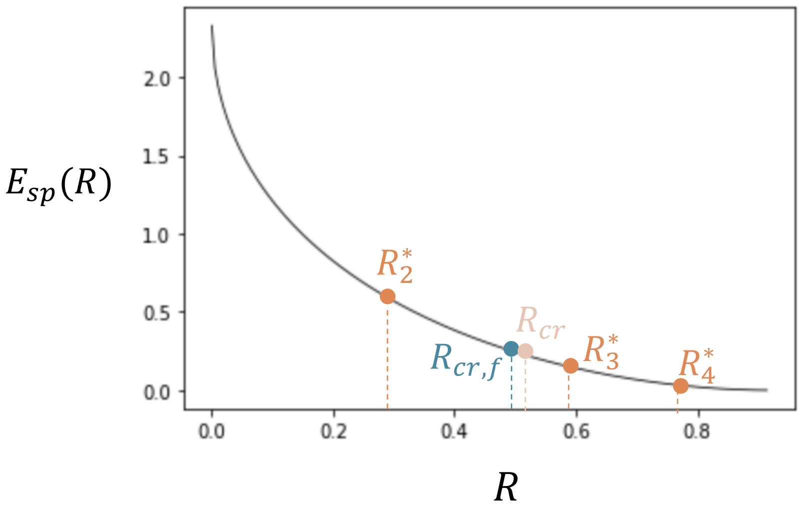
Sphere Packing Bound Below Critical Rate



$p = 0.2$



$p = 0.05$



$p = 0.01$

Conclusion

- We provide a practical implementation of the rubber coding method of Ahlswede *et al.* for feedback BSC.
- For the stochastic BSC with feedback, we show that our scheme is nearly optimal in a strong sense for certain parameters.
- A byproduct of the analysis is a strict enlargement of the rates for which the sphere-packing bound is known to be achievable with feedback BSC.

References

[1] E. R. Berlekamp, “Block coding for the binary symmetric channel with noiseless, delayless feedback,” in Error-Correcting Codes, H. B. Mann, Ed., the Mathematics Research Center, United States Army at the University of Wisconsin, Madison. Wiley New York, May 1968, pp. 61–68.

[2] K. Zigangirov, “On the number of correctable errors for transmission over a binary symmetrical channel with feedback,” Problemy Peredachi Informatsii, vol. 12, no. 2, pp. 3–19, 1976.

[3] J. Schalkwijk, “A class of simple and optimal strategies for block coding on the binary symmetric channel with noiseless feedback,” IEEE Trans. Inf. Theory, vol. 17, no. 3, pp. 283–287, 1971.

[4] R. Ahlswede, C. Deppe, and V. Lebedev, “Non-binary error correcting codes with noiseless feedback, localized errors, or both,” in Proc. IEEE Intl. Symp. on Inf. Theory (ISIT), 2006, pp. 2486–2487.

[5] M. Horstein, “Sequential transmission using noiseless feedback,” IEEE Trans. Inf. Theory, vol. 9, no. 3, pp. 136–143, 1963.

[6] C. T. Li and A. El Gamal, “An efficient feedback coding scheme with low error probability for discrete memoryless channels,”IEEE Transactions on Information Theory, vol. 61, no. 6, pp. 2953–2963,2015.