Efficient Reliable Communication in the Short Blocklength Regime Through List Decoding and Through Feedback

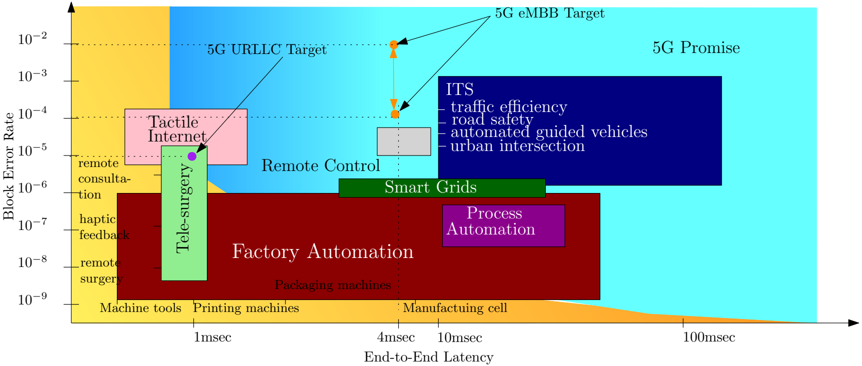
Hengjie Yang
Advisor: Richard Wesel

Ph.D. Defense
University of California, Los Angeles
May 16, 2022

Outline

1. CRC-Aided List Decoding of Convolutional Codes
2. Variable-Length Coding for Binary Channels With Full Feedback
3. Variable-Length Coding for Binary-Input Channels With Finite, Stop Feedback
4. Summary

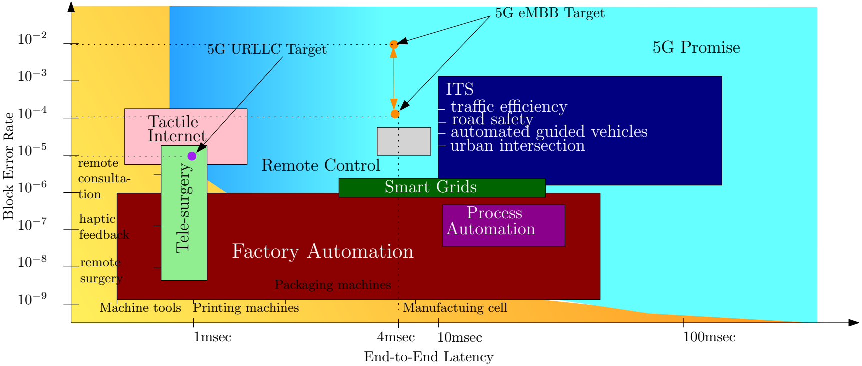
UCLA Samueli Electrical & Computer Engineering



M. Shirvanimoghaddam *et al.*, "Short blocklength codes for ultra-reliable low-latency communications," *IEEE Commun. Mag.*, Feb. 2019. Fig. 1.

UCLA Samueli Electrical & Computer Engineering

Ultra-Reliable Low-Latency Communication (URLLC) in 5G



These stringent requirements call for good short blocklength codes!

M. Shirvanimoghaddam *et al.*, "Short blocklength codes for ultra-reliable low-latency communications," *IEEE Commun. Mag.*, Feb. 2019. Fig. 1.

Define

- n: blocklength
- M: message size
- $\epsilon^*(n, M) \triangleq \inf\{\epsilon : \exists$ an (n, M, ϵ) fixed-length code$\}$

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, May 2010.

UCLA Samueli Electrical & Computer Engineering

Define

- n: blocklength
- M: message size
- $\epsilon^*(n, M) \triangleq \inf\{\epsilon : \exists$ an (n, M, ϵ) fixed-length code$\}$

Theorem 1 (Random-coding union (RCU) bound, Polyanskiy *et al.*, 2010)

Fix $n \in \mathbb{N}$, $M \in \mathbb{N}$, and a memoryless channel $(\mathcal{X}, \mathcal{Y}, W(Y|X))$.

$$\epsilon^*(n, M) \leq \mathbb{E}\Big[\min\left\{1, (M-1)\mathbb{P}\left[W^n(Y^n|\bar{X}^n) \geq W^n(Y^n|X^n)\right]\right\}\Big] \qquad (1)$$

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, May 2010.

Define

- n: blocklength
- M: message size
- $\epsilon^*(n, M) \triangleq \inf\{\epsilon : \exists$ an (n, M, ϵ) fixed-length code$\}$

Theorem 1 (Random-coding union (RCU) bound, Polyanskiy *et al.*, 2010)

Fix $n \in \mathbb{N}$, $M \in \mathbb{N}$, and a memoryless channel $(\mathcal{X}, \mathcal{Y}, W(Y|X))$.

$$\epsilon^*(n, M) \leq \mathbb{E}\left[\min\left\{1, (M-1)\mathbb{P}\left[W^n(Y^n|\bar{X}^n) \geq W^n(Y^n|X^n)\right]\right\}\right] \tag{1}$$

Theorem 2 (Meta-converse (MC) bound, Polyanskiy *et al.*, 2010)

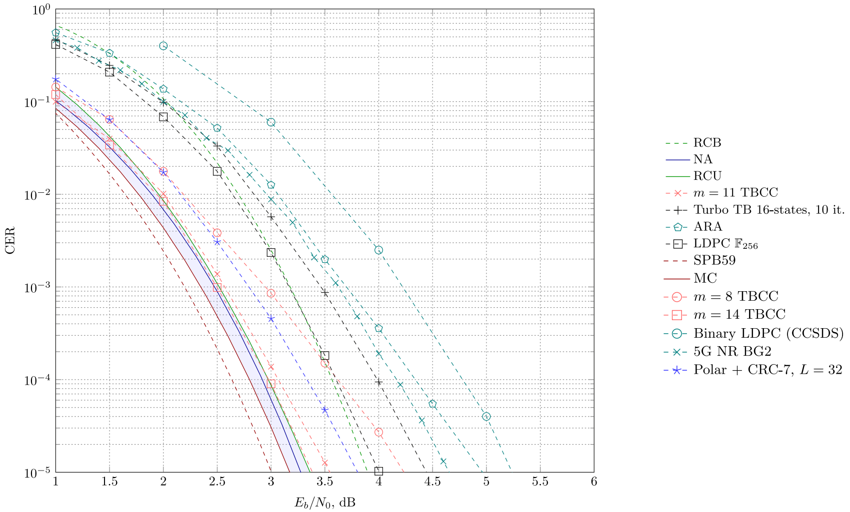
Fix $n \in \mathbb{N}$, $M \in \mathbb{N}$, and a memoryless channel $(\mathcal{X}, \mathcal{Y}, W(Y|X))$.

$$\epsilon^*(n, M) \geq \min_{P^n} \max_{Q^n} \left\{\alpha_{\frac{1}{M}}\left(P^n \times W^n, P^n \times Q^n\right)\right\} \tag{2}$$

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, May 2010.

UCLA Samueli Electrical & Computer Engineering

Contemporary Short Blocklength Code Performance



M. Coşkun *et al.*, "Efficient error-correcting codes in the short blocklength regime", *Physical Commun.*, Jun. 2019. Fig. 10. $n = 128$, $k = 64$, BI-AWGN channel.

UCLA Samueli Electrical & Computer Engineering

Can we approach the RCU bound at a reasonable complexity?

Based on...

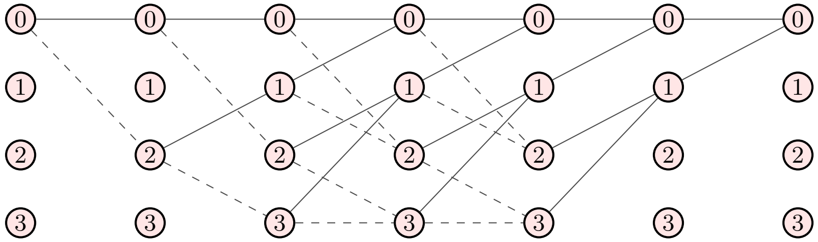
H. Yang, E. Liang, M. Pan, and R. D. Wesel, "CRC-Aided List Decoding of Convolutional Codes in the Short Blocklength Regime," *IEEE Trans. Inf. Theory*, Feb. 2022, early access.

UCLA Samueli Electrical & Computer Engineering

Serial List Viterbi Decoding (SLVD)

$\boldsymbol{c}$: $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$

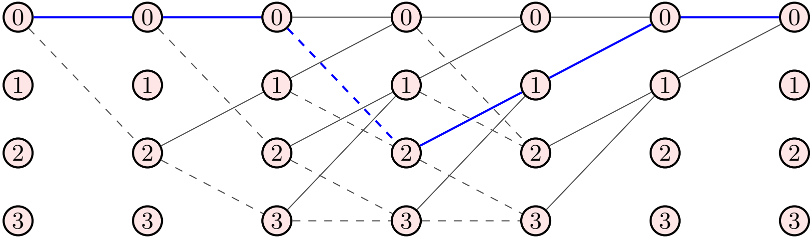
$\boldsymbol{y}$: $(2.1,-0.4)$ $(2.4,0.1)$ $(0.9,-0.5)$ $(1.9,0.1)$ $(-0.5,-1.8)$ $(1.1,0.1)$



Parameters setup: $k=2$, degree-2 CRC poly. $p(x)=x^2+1$, ZTCC $(5,7)$.

$\boldsymbol{c}$: $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$

$\boldsymbol{y}$: $(2.1,-0.4)$ $(2.4,0.1)$ $(0.9,-0.5)$ $(1.9,0.1)$ $(-0.5,-1.8)$ $(1.1,0.1)$



No.	Metric	Check?
1	13.53	No

$L = 1 \Leftarrow$

NACK! (If $\Psi = 1$)

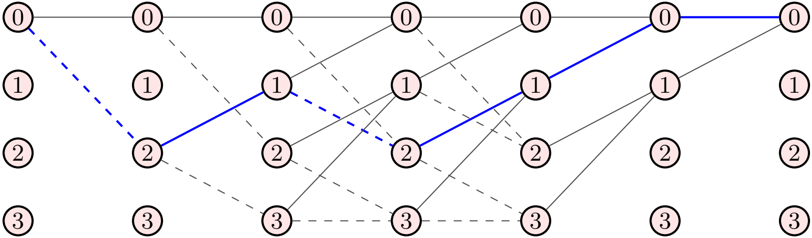
Parameters setup: $k = 2$, degree-2 CRC poly. $p(x) = x^2 + 1$, ZTCC $(5, 7)$.

UCLA Samueli Electrical & Computer Engineering

Serial List Viterbi Decoding (SLVD)

$\boldsymbol{c}$: $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$ $(0,0)$

$\boldsymbol{y}$: $(2.1,-0.4)$ $(2.4,0.1)$ $(0.9,-0.5)$ $(1.9,0.1)$ $(-0.5,-1.8)$ $(1.1,0.1)$

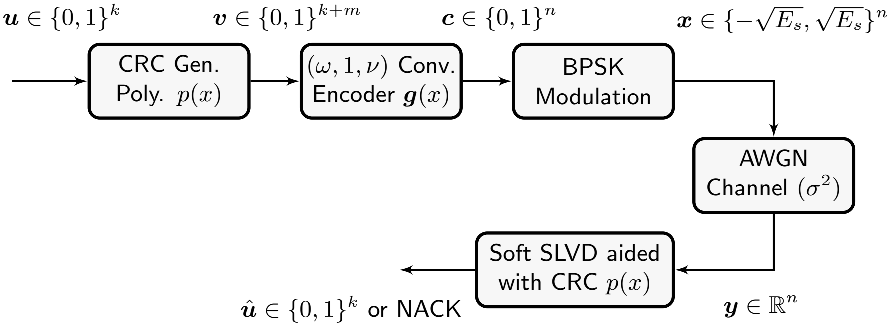


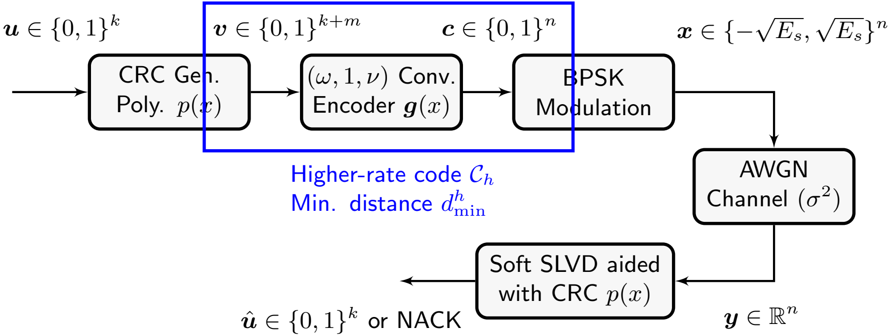
	No.	Metric	Check?
	1	13.53	No
$L=2 \Leftarrow$	2	15.13	Yes

Undetected Error (UE)! (If $\Psi \geq 2$)

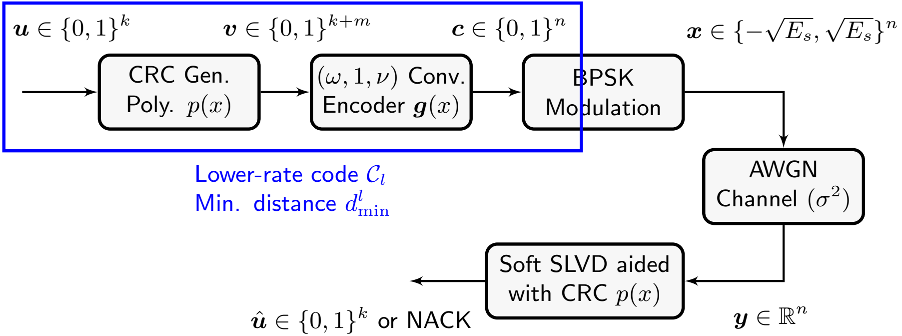
Parameters setup: $k=2$, degree-2 CRC poly. $p(x)=x^2+1$, ZTCC $(5,7)$.

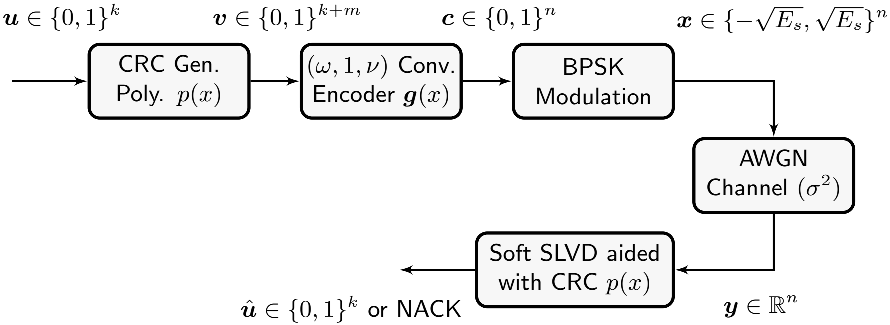
UCLA Samueli Electrical & Computer Engineering


$\boldsymbol{u} \in \{0,1\}^k$
$\boldsymbol{v} \in \{0,1\}^{k+m}$
$\boldsymbol{c} \in \{0,1\}^n$
$\boldsymbol{x} \in \{-\sqrt{E_s}, \sqrt{E_s}\}^n$
CRC Gen. Poly. $p(x)$
$(\omega, 1, \nu)$ Conv. Encoder $\boldsymbol{g}(x)$
BPSK Modulation
AWGN Channel (σ^2)
Soft SLVD aided with CRC $p(x)$
$\hat{\boldsymbol{u}} \in \{0,1\}^k$ or NACK
$\boldsymbol{y} \in \mathbb{R}^n$


$\boldsymbol{u} \in \{0,1\}^k$
$\boldsymbol{v} \in \{0,1\}^{k+m}$
$\boldsymbol{c} \in \{0,1\}^n$
$\boldsymbol{x} \in \{-\sqrt{E_s}, \sqrt{E_s}\}^n$
CRC Gen. Poly. $p(x)$
$(\omega, 1, \nu)$ Conv. Encoder $\boldsymbol{g}(x)$
BPSK Modulation
AWGN Channel (σ^2)
Higher-rate code $\mathcal{C}_h$
Min. distance $d_{\min}^h$
Soft SLVD aided with CRC $p(x)$
$\hat{\boldsymbol{u}} \in \{0,1\}^k$ or NACK
$\boldsymbol{y} \in \mathbb{R}^n$

CRC-Aided Convolutional Codes under SLVD





Two concatenated codes of interest:

- Zero-terminated convolutional codes (ZTCC) $\Rightarrow$ CRC-ZTCC
- Tail-biting convolutional codes (TBCC) $\Rightarrow$ CRC-TBCC

Supremum List Rank λ

$$\lambda \leq |\mathcal{C}_h| - |\mathcal{C}_l| + 1 = 2^{k+m} - 2^k + 1$$

$$\lambda \leq |\mathcal{C}_h| - |\mathcal{C}_l| + 1 = 2^{k+m} - 2^k + 1$$

Open problem: How to determine λ exactly for a given $\mathcal{C}_l$, $\mathcal{C}_h$?

$$\lambda \leq |\mathcal{C}_h| - |\mathcal{C}_l| + 1 = 2^{k+m} - 2^k + 1$$

Open problem: How to determine λ exactly for a given $\mathcal{C}_l$, $\mathcal{C}_h$?

$$1 \leq L \leq \min\{\lambda, \Psi\}$$

UCLA Samueli Electrical & Computer Engineering

Problem 1: Given a convolutional code, how to design the optimal CRC gen. poly. $p(x)$?

Problem 2: What is the performance-complexity trade-off of the resulting code?

UCLA Samueli Electrical & Computer Engineering

Problem: For a given ZTCC (or TBCC), CRC degree m, and SNR E_s/σ^2,

$$\min_{p(x)} \quad P_{e,\lambda} \tag{3}$$

where

$$p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2x^2 + a_1x + 1, \quad a_i \in \{0, 1\} \tag{4}$$

Problem: For a given ZTCC (or TBCC), CRC degree m, and SNR E_s/σ^2,

$$\min_{p(x)} \quad P_{e,\lambda} \tag{3}$$

where

$$p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2x^2 + a_1x + 1, \quad a_i \in \{0, 1\} \tag{4}$$

Obstacle: $P_{e,\lambda}$ does not admit an analytical expression.

Problem: For a given ZTCC (or TBCC), CRC degree m, and SNR E_s/σ^2,

$$\min_{p(x)} \quad P_{e,\lambda} \tag{3}$$

where

$$p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2x^2 + a_1x + 1, \quad a_i \in \{0,1\} \tag{4}$$

Obstacle: $P_{e,\lambda}$ does not admit an analytical expression.

Workaround: Take the union bound of $P_{e,\lambda}$ as the objective function!

$$P_{e,\lambda} \leq \sum_{\boldsymbol{c} \in \mathcal{C}_l \setminus \{\bar{\boldsymbol{c}}\}} \mathbb{P}\left(Z > \frac{1}{2} \|\boldsymbol{x}(\boldsymbol{c}) - \boldsymbol{x}(\bar{\boldsymbol{c}})\| \middle| \boldsymbol{X} = \boldsymbol{x}(\bar{\boldsymbol{c}}) \right) = \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}} \right) \tag{5}$$

where $C_{d^l_{\min}}, C_{d^l_{\min}+1}, \ldots, C_n$ denotes the distance spectrum of $\mathcal{C}_l$.

Degree-m DSO CRC polynomial at SNR $\sqrt{E_s/\sigma^2}$: defined as the solution to

$$\min_{p(x)} \quad \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right), \tag{6}$$

where

$$p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2 x^2 + a_1 x + 1, \quad \text{with } a_i \in \mathbb{F}_2 \tag{7}$$

Degree-m DSO CRC polynomial at SNR $\sqrt{E_s/\sigma^2}$: defined as the solution to

$$\min_{p(x)} \quad \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right), \tag{6}$$

where

$$p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2x^2 + a_1x + 1, \quad \text{with } a_i \in \mathbb{F}_2 \tag{7}$$

Theorem 3

For sufficiently large SNR E_s/σ^2, (6) simplifies to

$$\max_{p(x)} \quad d^l_{\min} \tag{8}$$

Theorem 4

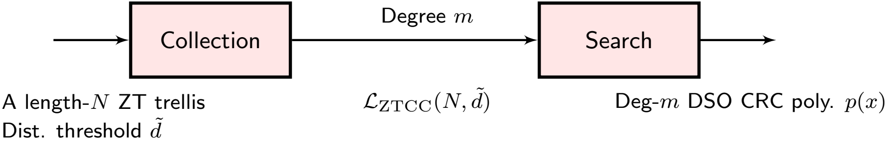
Fix CRC degree m and a higher-rate distance spectrum $B_{d^h_{\min}}, \ldots, B_n$. Define

$$w^* \triangleq \min \left\{ w \in \mathbb{N}_+ : \sum_{d=d^h_{\min}}^{w} B_d \geq 2^m \right\}. \quad (9)$$

For any degree-m CRC polynomial, we have $d^l_{\min} \leq 2w^*$.

DSO CRC Search Algorithm

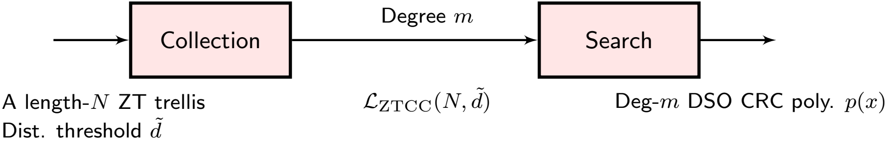
Lou et al.'s algorithm for ZTCC case



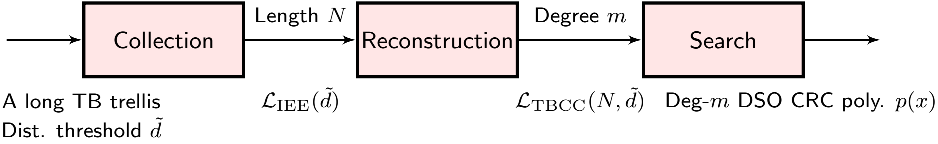
C. Lou *et al.*, "Convolutional-code-specific CRC code design," *IEEE Trans. Commun.*, Oct. 2015.

UCLA Samueli Electrical & Computer Engineering

Lou et al.'s algorithm for ZTCC case



Our algorithm for TBCC case



Contribution: A novel collection step that collects non-catastrophic IEEs.

C. Lou *et al.*, "Convolutional-code-specific CRC code design," *IEEE Trans. Commun.*, Oct. 2015.

UCLA Samueli Electrical & Computer Engineering

DSO CRC Polynomials for rate-$1/2$ ZTCC with $k = 64$

ν	ZTCC $\boldsymbol{g}(x)$	DSO CRC Polynomials $m=3$	4	5	6	7	8	9	10
3	$(13, 17)$	9	1B	2D	43	B5	107	313	50B
4	$(27, 31)$	F	15	33	4F	D3	13F	2AD	709
5	$(53, 75)$	9	11	25	49	EF	131	23F	73D
6	$(133, 171)$	F	1B	23	41	8F	113	2EF	629
7	$(247, 371)$	9	13	3F	5B	E9	17F	2A5	61D
8	$(561, 753)$	F	11	33	49	8B	19D	27B	4CF
9	$(1131, 1537)$	D	15	21	51	B7	1D5	20F	50D
10	$(2473, 3217)$	F	13	3D	5B	BB	105	20D	6BB

Parameter setup: $k = 64$. CRC polynomials are in hexadecimal. Optimum encoders are from [Lin and Costello].

Lin and Costello, "Error control coding", USA: Pearson, Table 12.1(c).

UCLA Samueli Electrical & Computer Engineering

ν	TBCC $\boldsymbol{g}(x)$	DSO CRC Polynomials							
		$m = 3$	4	5	6	7	8	9	10
3	$(13, 17)$	F	1F	2D	63	ED	107	349	49D
4	$(27, 31)$	F	11	33	4F	B5	1AB	265	4D1
5	$(53, 75)$	9	11	3F	63	BD	16D	349	41B
6	$(133, 171)$	F	1B	3D	7F	FF	145	2BD	571
7	$(247, 371)$	F	11	33	63	EF	145	3A1	5D7
8	$(561, 753)$	F	11	33	7F	FF	1AB	301	4F5
9	$(1131, 1537)$	D	15	33	51	C5	1FF	349	583
10	$(2473, 3217)$	F	1B	33	79	BB	199	217	4DD

Parameter setup: $k = 64$. CRC polynomials are in hexadecimal. Optimum encoders are from [Lin and Costello].

Lin and Costello, "Error control coding", USA: Pearson, Table 12.1(c).

UCLA Samueli Electrical & Computer Engineering

Performance measures:

- Prob. of correct decoding $P_{c,\Psi}$
- Prob. of UE $P_{e,\Psi}$
- Prob. of NACK $P_{\mathsf{NACK},\Psi}$

Performance measures:

- Prob. of correct decoding $P_{c,\Psi}$
- Prob. of UE $P_{e,\Psi}$
- Prob. of NACK $P_{\mathsf{NACK},\Psi}$

Question: How do these quantities vary with Ψ and SNR E_s/σ^2?

UCLA Samueli Electrical & Computer Engineering

Performance measures:

- Prob. of correct decoding $P_{c,\Psi}$
- Prob. of UE $P_{e,\Psi}$
- Prob. of NACK $P_{NACK,\Psi}$

Question: How do these quantities vary with Ψ and SNR E_s/σ^2?

Theorem 5

For a given CRC-aided convolutional code under SLVD at a fixed SNR, $P_{c,\Psi}$ and $P_{e,\Psi}$ are both **strictly** *increasing in Ψ, and will converge to $P_{c,\lambda}$ and $P_{e,\lambda}$, respectively, where $P_{c,\lambda} + P_{e,\lambda} = 1$.*

Implication: $\min_{\Psi}\left(P_{e,\Psi} + P_{NACK,\Psi}\right) = P_{e,\lambda}$ and $\Psi^* \geq \lambda$.

UCLA Samueli Electrical & Computer Engineering

Approximations to $P_{e,1}, P_{e,\lambda}, P_{\mathsf{NACK},1}$ as a Function of SNR

- Higher-rate distance spectrum: $B_{d^h_{\min}}, B_{d^h_{\min}+1}, \ldots, B_n$
- Lower-rate distance spectrum: $C_{d^l_{\min}}, C_{d^l_{\min}+1}, \ldots, C_n$

UCLA Samueli Electrical & Computer Engineering

- Higher-rate distance spectrum: $B_{d^h_{\min}}, B_{d^h_{\min}+1}, \ldots, B_n$
- Lower-rate distance spectrum: $C_{d^l_{\min}}, C_{d^l_{\min}+1}, \ldots, C_n$

$$P_{e,1} \leq \min\left\{2^{-m}, \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\} \approx \min\left\{2^{-m},\ C_{d^l_{\min}} Q\left(\sqrt{\frac{d^l_{\min}E_s}{\sigma^2}}\right)\right\}$$

UCLA Samueli Electrical & Computer Engineering

Approximations to $P_{e,1}, P_{e,\lambda}, P_{\text{NACK},1}$ as a Function of SNR

- Higher-rate distance spectrum: $B_{d^h_{\min}}, B_{d^h_{\min}+1}, \ldots, B_n$
- Lower-rate distance spectrum: $C_{d^l_{\min}}, C_{d^l_{\min}+1}, \ldots, C_n$

$$P_{e,1} \leq \min\left\{2^{-m}, \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\} \approx \min\left\{2^{-m},\ C_{d^l_{\min}} Q\left(\sqrt{\frac{d^l_{\min}E_s}{\sigma^2}}\right)\right\}$$

$$P_{e,\lambda} \leq \min\left\{1, \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\} \approx \min\left\{1, \sum_{d=d^l_{\min}}^{\tilde{d}} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\}$$

UCLA Samueli Electrical & Computer Engineering

Approximations to $P_{e,1}, P_{e,\lambda}, P_{\text{NACK},1}$ as a Function of SNR

- Higher-rate distance spectrum: $B_{d^h_{\min}}, B_{d^h_{\min}+1}, \ldots, B_n$
- Lower-rate distance spectrum: $C_{d^l_{\min}}, C_{d^l_{\min}+1}, \ldots, C_n$

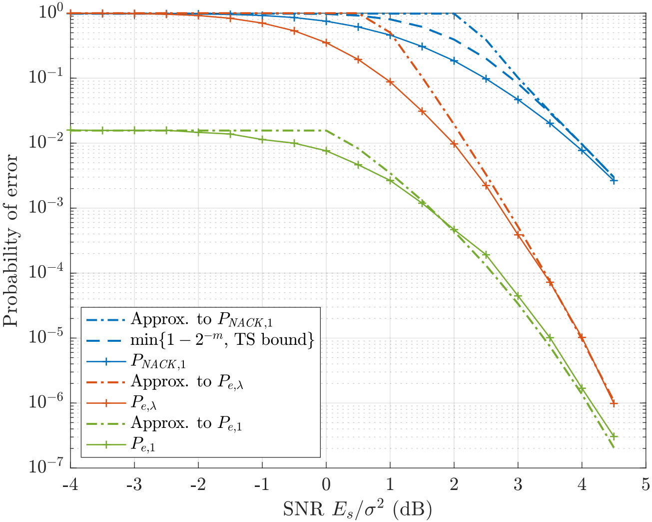
$$P_{e,1} \leq \min\left\{2^{-m}, \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\} \approx \min\left\{2^{-m},\ C_{d^l_{\min}} Q\left(\sqrt{\frac{d^l_{\min}E_s}{\sigma^2}}\right)\right\}$$

$$P_{e,\lambda} \leq \min\left\{1, \sum_{d=d^l_{\min}}^{n} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\} \approx \min\left\{1, \sum_{d=d^l_{\min}}^{\tilde{d}} C_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right)\right\}$$

$$P_{\mathit{NACK},1} \approx \min\left\{1-2^{-m}, \sum_{d=d^h_{\min}}^{\tilde{d}} B_d Q\left(\sqrt{\frac{dE_s}{\sigma^2}}\right) - C_{d^l_{\min}} Q\left(\sqrt{\frac{d^l_{\min}E_s}{\sigma^2}}\right)\right\}$$

UCLA Samueli Electrical & Computer Engineering

Example of Approximations



Setting: $k = 64$, degree-6 DSO CRC poly. 0x43, ZTCC $(13, 17)$, $\tilde{d} = 24$.
The tangential-sphere (TS) bound is also shown for $P_{NACK,1}$.

UCLA Samueli Electrical & Computer Engineering

Expected Terminating List Rank

Theorem 6

For a given CRC-aided convolutional code decoded with SLVD,

$$\lim_{\frac{E_s}{\sigma^2}\to 0} \mathbb{E}[L] = \mathbb{E}[L|\boldsymbol{X} = \boldsymbol{O}]. \tag{10}$$

UCLA Samueli Electrical & Computer Engineering

Theorem 6

For a given CRC-aided convolutional code decoded with SLVD,

$$\lim_{\frac{E_s}{\sigma^2} \to 0} \mathbb{E}[L] = \mathbb{E}[L|\boldsymbol{X} = \boldsymbol{O}]. \tag{10}$$

Lemma 1

There exists a possibly nonlinear lower-rate code $\mathcal{C}_l$ with

$$\mathbb{E}[L|\boldsymbol{X} = \boldsymbol{O}, \mathcal{C}_l] \leq 2^m.$$

UCLA Samueli Electrical & Computer Engineering

Theorem 7 (Parametric Approximation)

Define $\bar{L} \triangleq \mathbb{E}[L|\boldsymbol{X} = \boldsymbol{O}]$. *For a CRC-aided convolutional code with corresponding parameters* $\bar{L}$ *and* $P_{e,\lambda}$,

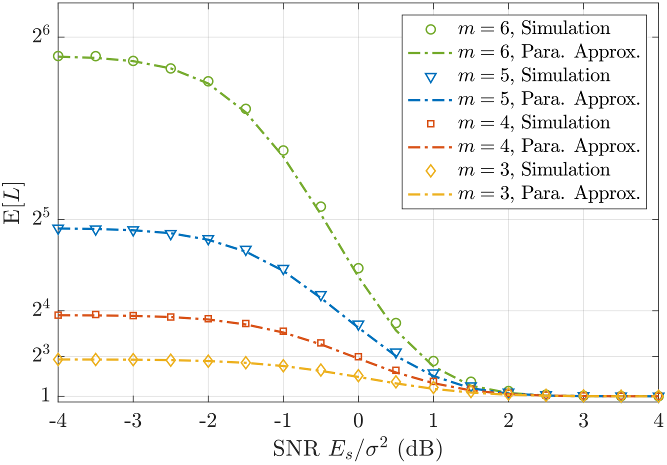
$$\mathbb{E}[L] \approx 1 - P_{e,\lambda} + P_{e,\lambda}\bar{L}. \tag{11}$$

Theorem 7 (Parametric Approximation)

Define $\bar{L} \triangleq \mathbb{E}[L|\boldsymbol{X} = \boldsymbol{O}]$*. For a CRC-aided convolutional code with corresponding parameters* $\bar{L}$ *and* $P_{e,\lambda}$*,*

$$\mathbb{E}[L] \approx 1 - P_{e,\lambda} + P_{e,\lambda}\bar{L}. \tag{11}$$

Remarks: Assume a target error prob. P_e^* and $\bar{L} \approx 2^m$ (true for CRC-ZTCC), $m \lessapprox -\log(P_e^*)$ is sufficient to maintain $\mathbb{E}[L] \leq 2$.



Setting: $k = 64$, degree-m DSO CRC polynomials for ZTCC $(13, 17)$.

UCLA Samueli Electrical & Computer Engineering

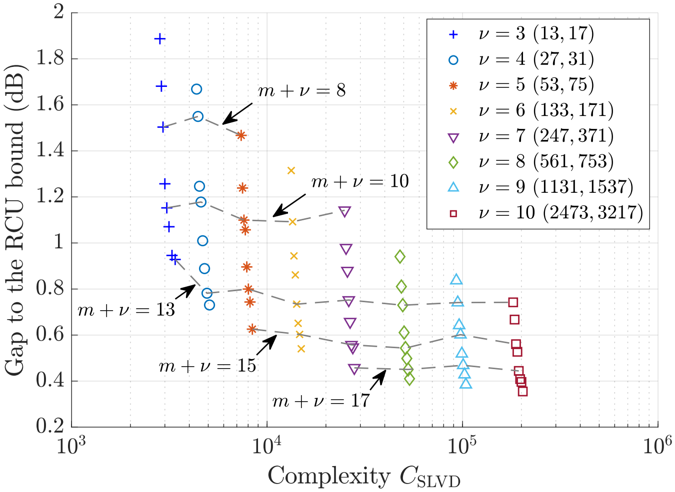
For our specific implementation of SLVD, three components comprise the average complexity of SLVD:

$$C_{\mathsf{SLVD}} = C_{\mathsf{SSV}} + C_{\mathsf{trace}} + C_{\mathsf{list}}. \tag{12}$$

Variable	CRC-ZTCC	CRC-TBCC
C_{SSV}	$(2^{\nu+1}-2)+1.5(2^{\nu+1}-2)+1.5(k+m-\nu)2^{\nu+1}+c_1[2(k+m+\nu)+1.5(k+m)]$	$1.5(k+m)2^{\nu+1}+2^{\nu}+3.5c_1(k+m)$
C_{trace}	$c_1(\mathbb{E}[L]-1)[2(k+m+\nu)+1.5(k+m)]$	$3.5c_1(\mathbb{E}[L]-1)(k+m)$
C_{list}	$c_2\mathbb{E}[I]\log(\mathbb{E}[I])$	
Notes	(i). c_1 and c_2 are computer-specific constants. (ii). $\mathbb{E}[I]$ denotes the average number of insertions (iii). For CRC-TBCC codes, $\mathbb{E}[I] \le (k+m)\mathbb{E}[L]+2^{\nu}-1$	

UCLA Samueli Electrical & Computer Engineering

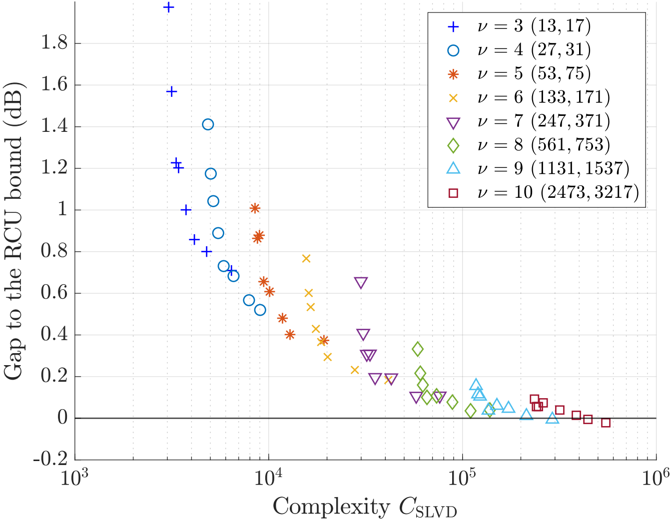
Performance-Complexity Trade-off for CRC-ZTCCs



Parameters setup: $k = 64$ and target error probability $P_{e,\lambda} = 10^{-4}$.

UCLA Samueli Electrical & Computer Engineering

Performance-Complexity Trade-off for CRC-TBCCs



Parameters setup: $k = 64$ and target error probability $P_{e,\lambda} = 10^{-4}$.

UCLA Samueli Electrical & Computer Engineering

UCLA Samueli Electrical & Computer Engineering

H. Yang and R. D. Wesel, "Finite-Blocklength Performance of Sequential Transmission Over BSC With Noiseless Feedback," in *Proc. IEEE Int. Sym. Inf. Theory (ISIT)*, Los Angeles, CA, USA, June 2020.

H. Yang, M. Pan, A. Antonini, and R. D. Wesel, "Sequential Transmission Over Binary Asymmetric Channels With Feedback," accepted for publication in the *IEEE Trans. Inf. Theory*, May 2022.

Feedback is Useful

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Feedback is Useful

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.

UCLA Samueli Electrical & Computer Engineering

Feedback is Useful

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.
- Achieve better error exponents: [Burnashev, 1976]. For any $R \in [0, C]$,

$$E(R) \triangleq \lim_{\epsilon \to 0} \frac{\log \frac{1}{\epsilon}}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right). \tag{13}$$

Feedback is Useful

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.
- Achieve better error exponents: [Burnashev, 1976]. For any $R \in [0, C]$,

$$E(R) \triangleq \lim_{\epsilon \to 0} \frac{\log \frac{1}{\epsilon}}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right). \tag{13}$$

- Achieve universality: LT codes (or fountain codes) [Luby, 2002]

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.
- Achieve better error exponents: [Burnashev, 1976]. For any $R \in [0, C]$,

$$E(R) \triangleq \lim_{\epsilon \to 0} \frac{\log \frac{1}{\epsilon}}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right). \tag{13}$$

- Achieve universality: LT codes (or fountain codes) [Luby, 2002]
- Improve first- and second-order coding rates: [Polyanskiy *et al.*, 2011] (for $\epsilon < 1/2$)

UCLA Samueli Electrical & Computer Engineering

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.
- Achieve better error exponents: [Burnashev, 1976]. For any $R \in [0, C]$,

$$E(R) \triangleq \lim_{\epsilon \to 0} \frac{\log \frac{1}{\epsilon}}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right). \qquad (13)$$

- Achieve universality: LT codes (or fountain codes) [Luby, 2002]
- Improve first- and second-order coding rates: [Polyanskiy *et al.*, 2011] (for $\epsilon < 1/2$)

Fixed-length transmission:

- Achieve better error exponents: e.g., Schalkwijk–Kailath scheme

UCLA Samueli Electrical & Computer Engineering

Feedback is Useful

Bad news: Feedback does not improve the capacity of a memoryless channel [Shannon, 1956].

Variable-length transmission:

- Simplify coding schemes: e.g., Horstein scheme, posterior matching.
- Achieve better error exponents: [Burnashev, 1976]. For any $R \in [0, C]$,

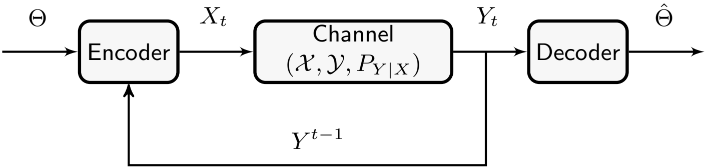
$$E(R) \triangleq \lim_{\epsilon \to 0} \frac{\log \frac{1}{\epsilon}}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right). \tag{13}$$

- Achieve universality: LT codes (or fountain codes) [Luby, 2002]
- Improve first- and second-order coding rates: [Polyanskiy *et al.*, 2011] (for $\epsilon < 1/2$)

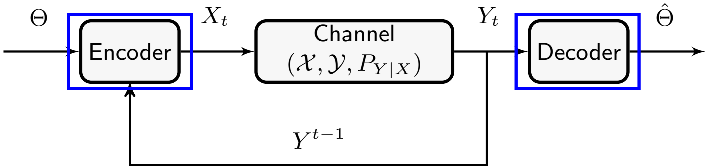
Fixed-length transmission:

- Achieve better error exponents: e.g., Schalkwijk–Kailath scheme
- Improve second-order coding rate for compound-dispersion DMCs: [Wagner, 2020].

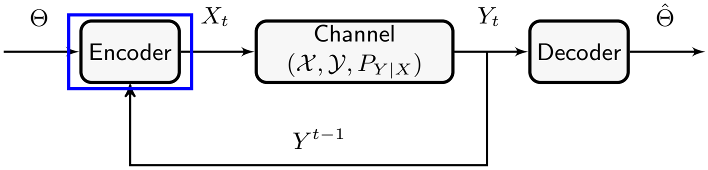
UCLA Samueli Electrical & Computer Engineering



Given $M \in \mathbb{N}_+$, $l > 0$, $\epsilon \in (0,1)$, we want to specify an (l, M, ϵ) variable-length feedback (VLF) code.



Codebook $U \in \mathcal{U}$: can be designed "on the fly" with full feedback



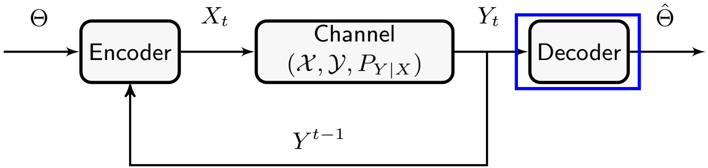
$[M] \triangleq \{1, 2, \ldots, M\}$.

Encoding function $e_t : \mathcal{U} \times [M] \times \mathcal{Y}^{t-1} \to \mathcal{X}$:

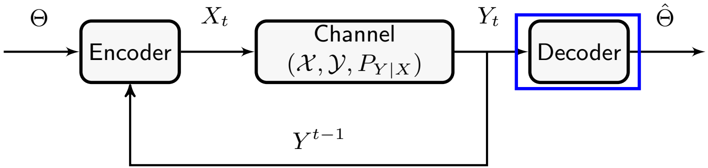
$$X_t = e_t(U, \Theta, Y^{t-1}), \quad t \in \mathbb{N}_+$$

where $\Theta \sim \mathrm{Unif}\,([M])$.

UCLA Samueli Electrical & Computer Engineering



Decoding function $g_t : \mathcal{U} \times \mathcal{Y}^t \to [M]$: providing the best estimate of Θ at time t, $t \in \mathbb{N}_+$.

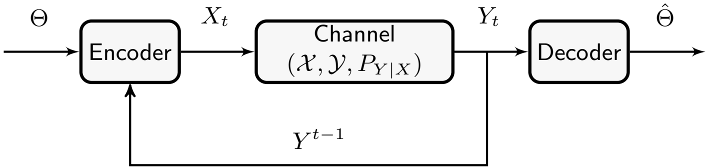


Stopping time $\tau \in \mathbb{N}$: a function of filtration $\mathcal{F}_t = \sigma\{U, Y^t\}$ and must satisfy $\mathbb{E}[\tau] \leq l$.

Final decision: $\hat{\Theta} = g_\tau(Y^\tau)$

τ also needs to satisfy

$$P_e \triangleq \mathbb{P}[\Theta \neq \hat{\Theta}] \leq \epsilon.$$



Goal: Determine $l^*(M, \epsilon) \triangleq \min\{l : \exists (l, M, \epsilon)$ VLF code$\}$

UCLA Samueli Electrical & Computer Engineering

Naghshvar et al.'s SED Coding Scheme

For symmetric binary-input channels, Naghshvar *et al.* constructed a particular deterministic VLF code.

M. Naghshvar *et al.*, "Optimal reliability over a class of binary-input channels with feedback," *IEEE Inf. Theory Workshop*, Sep. 2012.

For symmetric binary-input channels, Naghshvar *et al.* constructed a particular deterministic VLF code.

Belief state vector $\boldsymbol{\rho}(t)$:

$$\boldsymbol{\rho}(t) \triangleq \begin{bmatrix} \rho_1(t) & \rho_2(t) & \cdots & \rho_M(t) \end{bmatrix} \quad (14)$$

where $\rho_i(t) \triangleq \mathbb{P}[\Theta = i|Y^t]$, $i \in [M]$, $t \in \mathbb{N}$. By default, $\rho_i(0) = 1/M$, $i \in [M]$.

M. Naghshvar *et al.*, "Optimal reliability over a class of binary-input channels with feedback," *IEEE Inf. Theory Workshop*, Sep. 2012.

For symmetric binary-input channels, Naghshvar *et al.* constructed a particular deterministic VLF code.

Belief state vector $\boldsymbol{\rho}(t)$:

$$\boldsymbol{\rho}(t) \triangleq \begin{bmatrix} \rho_1(t) & \rho_2(t) & \cdots & \rho_M(t) \end{bmatrix} \quad (14)$$

where $\rho_i(t) \triangleq \mathbb{P}[\Theta = i | Y^t]$, $i \in [M]$, $t \in \mathbb{N}$. By default, $\rho_i(0) = 1/M$, $i \in [M]$.

Bayes' update of $\boldsymbol{\rho}(t)$: Upon receiving $Y_{t+1} = y_{t+1}$,

$$\rho_i(t+1) = \frac{\rho_i(t) P_{Y|X}(y_{t+1}|x_{t+1,i})}{\sum_{j \in [M]} \rho_j(t) P_{Y|X}(y_{t+1}|x_{t+1,j})}, \quad i \in [M], \quad (15)$$

where $x_{t+1,i} \in \mathcal{X}$ denotes the input symbol for $i \in [M]$ at time $t+1$.

M. Naghshvar *et al.*, "Optimal reliability over a class of binary-input channels with feedback," *IEEE Inf. Theory Workshop*, Sep. 2012.

The SED bipartition (defining codebook U) : Let $S_0(t)$ and $S_1(t)$ be a bipartition of $[M]$. Define

$$\pi_x(t) \triangleq \sum_{i\in S_x(t)} \rho_i(t), \quad x \in \{0,1\}. \tag{16}$$

At time $t+1$, upon obtaining $\boldsymbol{\rho}(t)$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$0 \le \pi_0(t) - \pi_1(t) \le \min_{i\in S_0(t)} \rho_i(t). \tag{17}$$

UCLA Samueli Electrical & Computer Engineering

The SED bipartition (defining codebook U) : Let $S_0(t)$ and $S_1(t)$ be a bipartition of $[M]$. Define

$$\pi_x(t) \triangleq \sum_{i\in S_x(t)} \rho_i(t), \quad x \in \{0,1\}. \tag{16}$$

At time $t+1$, upon obtaining $\boldsymbol{\rho}(t)$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$0 \le \pi_0(t) - \pi_1(t) \le \min_{i\in S_0(t)} \rho_i(t). \tag{17}$$

Encoder: After SED bipartition of $[M]$ into $S_0(t), S_1(t)$,

$$X_{t+1} = \begin{cases} 0, & \text{if } \Theta \in S_0(t) \\ 1, & \text{if } \Theta \in S_1(t) \end{cases} \tag{18}$$

UCLA Samueli Electrical & Computer Engineering

The SED bipartition (defining codebook U) : Let $S_0(t)$ and $S_1(t)$ be a bipartition of $[M]$. Define

$$\pi_x(t) \triangleq \sum_{i \in S_x(t)} \rho_i(t), \quad x \in \{0,1\}. \tag{16}$$

At time $t+1$, upon obtaining $\boldsymbol{\rho}(t)$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$0 \leq \pi_0(t) - \pi_1(t) \leq \min_{i \in S_0(t)} \rho_i(t). \tag{17}$$

Encoder: After SED bipartition of $[M]$ into $S_0(t), S_1(t)$,

$$X_{t+1} = \begin{cases} 0, & \text{if } \Theta \in S_0(t) \\ 1, & \text{if } \Theta \in S_1(t) \end{cases} \tag{18}$$

- A straightforward algo.: $O(k2^k)$

UCLA Samueli Electrical & Computer Engineering

The SED bipartition (defining codebook U) : Let $S_0(t)$ and $S_1(t)$ be a bipartition of $[M]$. Define

$$\pi_x(t) \triangleq \sum_{i \in S_x(t)} \rho_i(t), \quad x \in \{0,1\}. \tag{16}$$

At time $t+1$, upon obtaining $\boldsymbol{\rho}(t)$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$0 \le \pi_0(t) - \pi_1(t) \le \min_{i \in S_0(t)} \rho_i(t). \tag{17}$$

Encoder: After SED bipartition of $[M]$ into $S_0(t), S_1(t)$,

$$X_{t+1} = \begin{cases} 0, & \text{if } \Theta \in S_0(t) \\ 1, & \text{if } \Theta \in S_1(t) \end{cases} \tag{18}$$

- A straightforward algo.: $O(k2^k)$
- A type-based algo. for relaxed SED bipartition [Antonini *et al.*, 2020]: $O(k^2)$

A. Antonini *et al.*, "Low complexity algorithms for transmission of short blocks over the BSC with full feedback," *IEEE Int. Sym. Inf. Theory (ISIT)*, Jun. 2020.

UCLA Samueli Electrical & Computer Engineering

Decoder: Upon receiving y_{t+1}, the decoder obtains $\boldsymbol{\rho}(t+1)$ with Bayes' rule. The decoder adopts

$$\tau \triangleq \min\left\{t \in \mathbb{N} : \max_{i\in[M]} \rho_i(t) \geq 1-\epsilon\right\}. \tag{19}$$

$$\hat{\Theta} \triangleq \arg\max_{i\in[M]} \rho_i(\tau). \tag{20}$$

Decoder: Upon receiving y_{t+1}, the decoder obtains $\boldsymbol{\rho}(t+1)$ with Bayes' rule. The decoder adopts

$$\tau \triangleq \min\left\{t \in \mathbb{N} : \max_{i\in[M]} \rho_i(t) \geq 1-\epsilon\right\}. \tag{19}$$

$$\hat{\Theta} \triangleq \arg\max_{i\in[M]} \rho_i(\tau). \tag{20}$$

Error probability: $P_e = \mathbb{E}\left[1 - \max_{i\in[M]} \rho_i(\tau)\right] \leq \epsilon$.

UCLA Samueli Electrical & Computer Engineering

Why is the SED coding rule interesting?

Log-likelihood ratio $U_j(t)$:

$$U_j(t) \triangleq \log \frac{\rho_j(t)}{1-\rho_j(t)}, \quad j \in [M] \tag{21}$$

Why is the SED coding rule interesting?

Log-likelihood ratio $U_j(t)$:

$$U_j(t) \triangleq \log \frac{\rho_j(t)}{1-\rho_j(t)}, \quad j \in [M] \tag{21}$$

Theorem 8 (Naghshvar *et al.*, 2012)

For symmetric binary-input channels, the SED bipartition yields a two-stage submartingale $\{U_i(t)\}_{t=0}^{\infty}$*:*

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] \geq U_i(t) + C, \quad \text{if } U_i(t) < 0 \tag{22a}$$

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] = U_i(t) + C_1 \quad \text{if } U_i(t) \geq 0 \tag{22b}$$

$$|U_i(t+1) - U_i(t)| \leq C_2, \tag{22c}$$

where $C = \max_{P_X} I(X;Y)$,

$$C_1 \triangleq \max_{x,x' \in \mathcal{X}} D\left(P_{Y|X=x} \| P_{Y|X=x'}\right) \tag{23}$$

$$C_2 \triangleq \max_{y \in \mathcal{Y}} \log \frac{\max_{x \in \mathcal{X}} P_{Y|X}(y|x)}{\min_{x \in \mathcal{X}} P_{Y|X}(y|x)}. \tag{24}$$

UCLA Samueli Electrical & Computer Engineering

Log-likelihood ratio $U_j(t)$:

$$U_j(t) \triangleq \log \frac{\rho_j(t)}{1-\rho_j(t)}, \quad j \in [M] \tag{21}$$

Theorem 8 (Naghshvar *et al.*, 2012)

For symmetric binary-input channels, the SED bipartition yields a two-stage submartingale $\{U_i(t)\}_{t=0}^{\infty}$:

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] \geq U_i(t) + C, \quad \textit{if } U_i(t) < 0 \tag{22a}$$

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] = U_i(t) + C_1 \quad \textit{if } U_i(t) \geq 0 \tag{22b}$$

$$|U_i(t+1) - U_i(t)| \leq C_2, \tag{22c}$$

where $C = \max_{P_X} I(X;Y)$,

$$C_1 \triangleq \max_{x,x' \in \mathcal{X}} D\left(P_{Y|X=x} \| P_{Y|X=x'}\right) \tag{23}$$

$$C_2 \triangleq \max_{y \in \mathcal{Y}} \log \frac{\max_{x\in\mathcal{X}} P_{Y|X}(y|x)}{\min_{x\in\mathcal{X}} P_{Y|X}(y|x)}. \tag{24}$$

For typical DMC, $0 < C \leq C_1 \leq C_2 < \infty$.

UCLA Samueli Electrical & Computer Engineering

Naghshvar et al.'s Bound for Binary Symmetric Channel (BSC)

Theorem 9 (Naghshvar *et al.*, 2015)

Fix $M \in \mathbb{N}_+$ *and* $\epsilon \in (0, 1/2)$*. The* (l, M, ϵ) *VLF code constructed from the SED coding rule for BSC satisfies*

$$l \leq \frac{\log M + \log\log\frac{M}{\epsilon}}{C} + \frac{\log\frac{1}{\epsilon} + 1}{C_1} + \frac{96 \cdot 2^{2C_2}}{CC_1} \quad (25)$$

M. Naghshvar *et al.*, "Extrinsic Jensen-Shannon divergence: applications to variable-length coding," *IEEE Trans. Inf. Theory*, Apr. 2015.

Theorem 9 (Naghshvar *et al.*, 2015)

Fix $M \in \mathbb{N}_+$ and $\epsilon \in (0, 1/2)$. The (l, M, ϵ) VLF code constructed from the SED coding rule for BSC satisfies

$$l \leq \frac{\log M + \log\log\frac{M}{\epsilon}}{C} + \frac{\log\frac{1}{\epsilon} + 1}{C_1} + \frac{96 \cdot 2^{2C_2}}{CC_1} \tag{25}$$

M. Naghshvar *et al.*, "Extrinsic Jensen-Shannon divergence: applications to variable-length coding," *IEEE Trans. Inf. Theory*, Apr. 2015.

Theorem 10 (Polyanskiy *et al.*, 2011)

Fix $M \in \mathbb{N}_+$ and $\epsilon \in (0, 1/2)$. There exists an (l, M, ϵ) variable-length stop-feedback (VLSF) code for DMC with bounded information density with

$$l \leq \frac{\log(M-1)}{C} + \frac{\log\frac{1}{\epsilon}}{C} + \frac{a_0}{C} \tag{26}$$

where $a_0 \triangleq \sup_{x\in\mathcal{X}, y\in\mathcal{Y}} \iota(x; y)$.

Y. Polyanskiy *et al.*, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, Aug. 2011.

Issue: Even with full feedback, Naghshvar's result is much looser than Polyanskiy's.

Issue: Even with full feedback, Naghshvar's result is much looser than Polyanskiy's.

Numerical example: For BSC(0.11) with $C = 0.5$ and $\epsilon = 10^{-3}$,

$$l \leq \frac{\log M + \log\log M}{0.5} + 5352.67 \quad \text{(Naghshvar's bound for VLF code)} \tag{27}$$

$$l \leq \frac{\log(M-1)}{0.5} + 21.93 \quad \text{(Polyanskiy's bound for VLSF code)} \tag{28}$$

UCLA Samueli Electrical & Computer Engineering

Issue: Even with full feedback, Naghshvar's result is much looser than Polyanskiy's.

Numerical example: For BSC(0.11) with $C = 0.5$ and $\epsilon = 10^{-3}$,

$$l \leq \frac{\log M + \log\log M}{0.5} + 5352.67 \quad \text{(Naghshvar's bound for VLF code)} \tag{27}$$

$$l \leq \frac{\log(M-1)}{0.5} + 21.93 \quad \text{(Polyanskiy's bound for VLSF code)} \tag{28}$$

Our goal: Seek a new upper bound for SED coding rule better than Polyanskiy's bound!

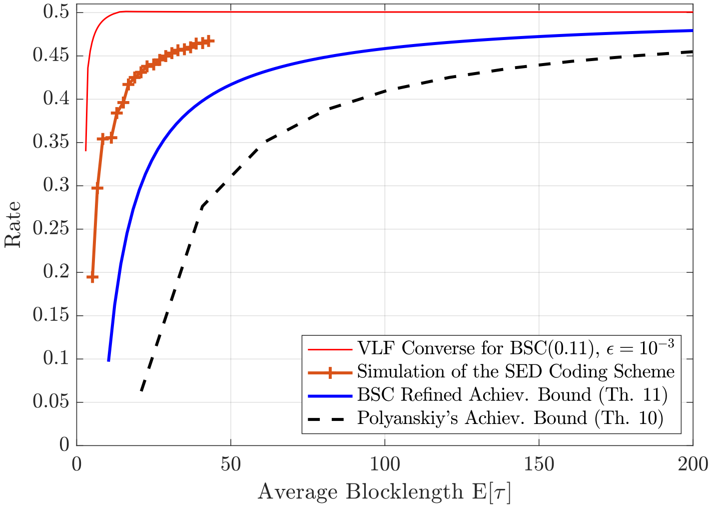
Our Main Result for BSC

Theorem 11

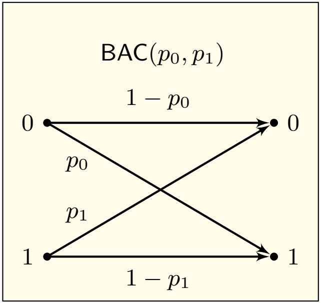
Fix M and $\epsilon \in (0,1/2)$. The (l, M, ϵ) VLF code constructed from the SED coding scheme for BSC(p), $p \in (0,1/2)$, satisfies

$$l < \frac{\log M + \frac{1}{q}\log 2q}{C} + \frac{\log \frac{1-\epsilon}{\epsilon} + C_2}{C_1} + 2^{-C_2}C_2\left(\frac{1}{C} - \frac{1}{C_1} + \frac{\frac{1}{q}\log 2q}{CC_2}\right)\frac{1 - \frac{\epsilon}{1-\epsilon}2^{-C_2}}{1-2^{-C_2}}. \tag{29}$$

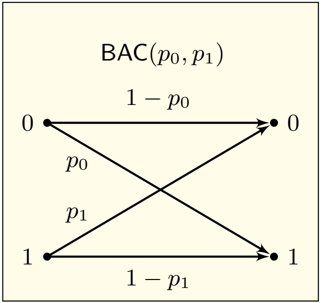
Numerical Evaluation for BSC(0.11) and $\epsilon = 10^{-3}$



UCLA Samueli Electrical & Computer Engineering

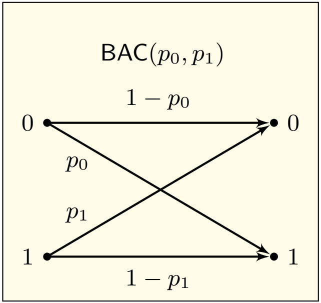

BAC(p_0, p_1)
$1-p_0$
0
0
p_0
p_1
1
1
$1-p_1$

Binary Asymmetric Channel (BAC)



Special cases:

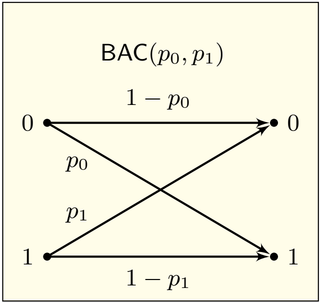
- Regularized BAC: $p_0 \in (0, 1/2)$ and $p_0 \leq p_1 \leq 1 - p_0$.



Special cases:

- Regularized BAC: $p_0 \in (0, 1/2)$ and $p_0 \le p_1 \le 1 - p_0$.
- BSC: $p_0 = p_1 \in (0, 1/2)$.

Binary Asymmetric Channel (BAC)



Special cases:

- Regularized BAC: $p_0 \in (0, 1/2)$ and $p_0 \leq p_1 \leq 1 - p_0$.
- BSC: $p_0 = p_1 \in (0, 1/2)$.

Every BAC can be transformed into an equivalent regularized BAC.

Fact 1: Consider a BAC(p_0, p_1). Define $z \triangleq 2^{\frac{h(p_0)-h(p_1)}{1-p_0-p_1}}$. Then,

$$C = \frac{p_0 h(p_1)}{1-p_0-p_1} - \frac{(1-p_1)h(p_0)}{1-p_0-p_1} + \log(1+z), \tag{30}$$

$$\pi_0^* = \frac{1-p_1(1+z)}{(1-p_0-p_1)(1+z)}, \tag{31}$$

$$\pi_1^* = \frac{(1-p_0)(1+z)-1}{(1-p_0-p_1)(1+z)}. \tag{32}$$

Further ,for regularized BAC(p_0, p_1), $0 < \pi_1^* \leq \pi_0^* < 1$.

UCLA Samueli Electrical & Computer Engineering

Fact 1: Consider a BAC(p_0, p_1). Define $z \triangleq 2^{\frac{h(p_0)-h(p_1)}{1-p_0-p_1}}$. Then,

$$C = \frac{p_0 h(p_1)}{1-p_0-p_1} - \frac{(1-p_1)h(p_0)}{1-p_0-p_1} + \log(1+z), \tag{30}$$

$$\pi_0^* = \frac{1-p_1(1+z)}{(1-p_0-p_1)(1+z)}, \tag{31}$$

$$\pi_1^* = \frac{(1-p_0)(1+z)-1}{(1-p_0-p_1)(1+z)}. \tag{32}$$

Further ,for regularized BAC(p_0, p_1), $0 < \pi_1^* \leq \pi_0^* < 1$.

Fact 2: For a regularized BAC(p_0, p_1),

$$C_1 = D\Big(P_{Y|X=1} \| P_{Y|X=0}\Big), \tag{33}$$

$$C_2 = \log \frac{P_{Y|X}(1|1)}{P_{Y|X}(1|0)} = \log \frac{1-p_1}{p_0}. \tag{34}$$

Posterior matching principle: we want to have

$$[\pi_0(t), \pi_1(t)] \approx [\pi_0^*, \pi_1^*], \quad \forall t \in \mathbb{N} \tag{35}$$

Posterior matching principle: we want to have

$$[\pi_0(t), \pi_1(t)] \approx [\pi_0^*, \pi_1^*], \quad \forall t \in \mathbb{N} \tag{35}$$

Generalized SED bipartition: At time $t+1$, let $\hat{i} \triangleq \arg\max_{j\in[M]} \rho_j(t)$.

- If $\rho_{\hat{i}}(t) < \pi_1^*$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$-\frac{\min_{i\in S_1(t)} \rho_i(t)}{\pi_1^*} \le \frac{\pi_0(t)}{\pi_0^*} - \frac{\pi_1(t)}{\pi_1^*} \le \frac{\min_{i\in S_0(t)} \rho_i(t)}{\pi_0^*}. \tag{36}$$

- If $\rho_{\hat{i}}(t) \ge \pi_1^*$, exclusively assign $S_0(t) = [M] \setminus \{\hat{i}\}$, $S_1(t) = \{\hat{i}\}$.

UCLA Samueli Electrical & Computer Engineering

Posterior matching principle: we want to have

$$[\pi_0(t), \pi_1(t)] \approx [\pi_0^*, \pi_1^*], \quad \forall t \in \mathbb{N} \tag{35}$$

Generalized SED bipartition: At time $t+1$, let $\hat{i} \triangleq \arg\max_{j\in[M]} \rho_j(t)$.

- If $\rho_{\hat{i}}(t) < \pi_1^*$, partition $[M]$ into $S_0(t)$ and $S_1(t)$ s.t.

$$-\frac{\min_{i\in S_1(t)} \rho_i(t)}{\pi_1^*} \le \frac{\pi_0(t)}{\pi_0^*} - \frac{\pi_1(t)}{\pi_1^*} \le \frac{\min_{i\in S_0(t)} \rho_i(t)}{\pi_0^*}. \tag{36}$$

- If $\rho_{\hat{i}}(t) \ge \pi_1^*$, exclusively assign $S_0(t) = [M] \setminus \{\hat{i}\}$, $S_1(t) = \{\hat{i}\}$.

Encoder & Decoder: same as the BSC case

UCLA Samueli Electrical & Computer Engineering

Lemma 2

Fix a regularized BAC(p_0, p_1). The generalized SED coding rule induces the following submartingale

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] \geq U_i(t) + C, \quad \textit{if } U_i(t) < 0 \tag{37a}$$

$$\mathbb{E}[U_i(t+1)|\mathcal{F}_t, \Theta = i] \geq U_i(t) + C_1 \quad \textit{if } U_i(t) \geq 0 \tag{37b}$$

$$|U_i(t+1) - U_i(t)| \leq C_2 \tag{37c}$$

UCLA Samueli Electrical & Computer Engineering

Theorem 12

Fix a regularized BAC(p_0, p_1). *Let* $\lambda \triangleq \pi_1^*/\pi_0^* \in (0,1]$. *For* $\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_M]$ *satisfying* $\max_{i\in[M]} \rho_i < \pi_1^*$, *define function* $f : 2^{[M]} \to \mathbb{R}$:

$$f(S) \triangleq \lambda\big(\pi_1(S) - \lambda\pi_0(S)\big)\mathbf{1}_{\{\pi_1(S)\geq\lambda\pi_0(S)\}} + \big(\lambda\pi_0(S) - \pi_1(S)\big)\mathbf{1}_{\{\pi_1(S)<\lambda\pi_0(S)\}}, \quad (38)$$

where

$$\pi_0(S) \triangleq \sum_{i\in S} \rho_i, \quad \pi_1(S) \triangleq \sum_{i\in[M]\setminus S} \rho_i \quad (39)$$

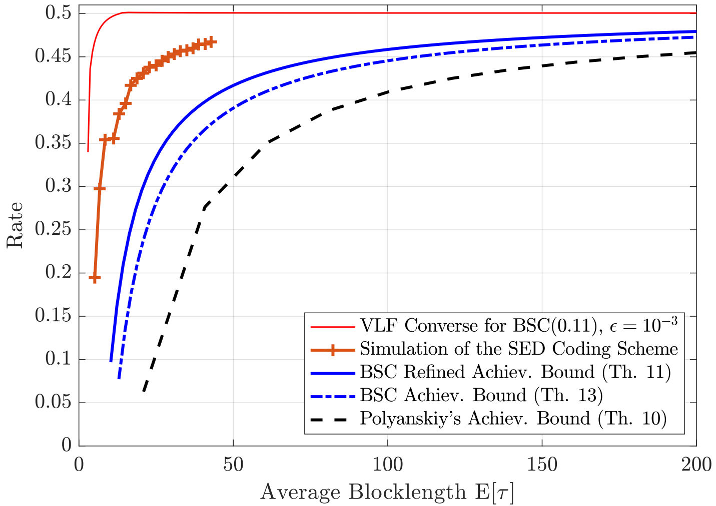
If $S_0^* \subseteq [M]$ *minimizes* (38), *then, the partition* $(S_0^*, [M]\setminus S_0^*)$ *satisfies* (36).

Theorem 13

Fix $M \in \mathbb{N}_+$ *and* $\epsilon \in (0, 1/2)$. *The* (l, M, ϵ) *VLF code constructed from the generalized SED coding scheme for regularized BAC*(p_0, p_1) *satisfies*

$$l < \frac{\log M}{C} + \frac{\log \frac{1-\epsilon}{\epsilon} + C_2}{C_1} + C_2 \left(\frac{1}{C} - \frac{1}{C_1} \right) \frac{1 - \frac{\epsilon}{1-\epsilon} 2^{-C_2}}{1 - 2^{-C_2}}. \tag{40}$$

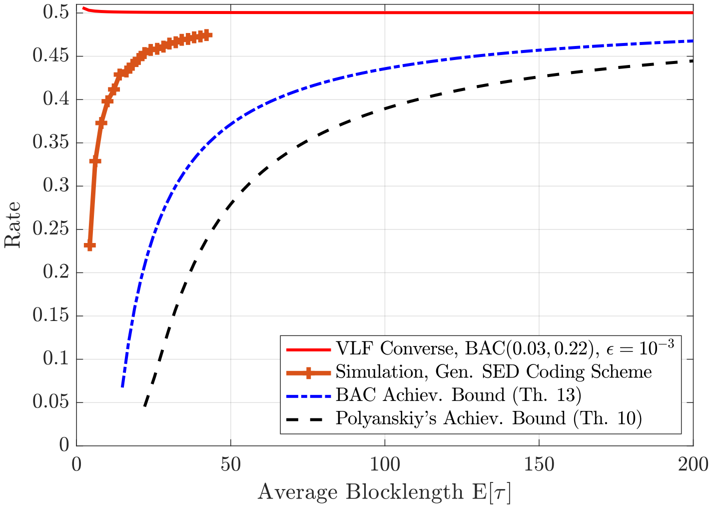
Numerical Evaluation for BSC(0.11) and $\epsilon = 10^{-3}$



Parameters: $C = 0.5$, $C_1 = 2.3527$, $C_2 = 3.02$.

UCLA Samueli Electrical & Computer Engineering

Numerical Evaluation for BAC$(0.03, 0.22)$ and $\epsilon = 10^{-3}$



Parameters: $C = 0.5$, $C_1 = 3.1954$, $C_2 = 4.7$.

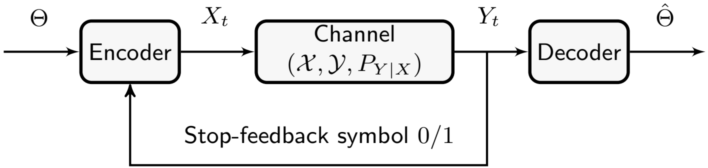
UCLA Samueli Electrical & Computer Engineering

Based on ...

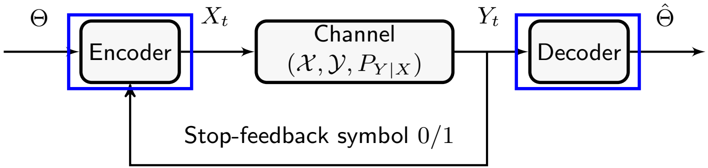
H. Yang, R. C. Yavas, V. Kostina, and R. D. Wesel, "Variable-Length Stop-Feedback Codes With Finite Optimal Decoding Times for BI-AWGN Channels," accepted for presentation at *IEEE Int. Sym. Inf. Theory (ISIT)*, Espoo, Finland, June 2022.

H. Yang, R. C. Yavas, V. Kostina, and R. D. Wesel, "Variable-Length Coding for Binary-Input Channels With Finite Stop Feedback," to be submitted to *IEEE Trans. Inf. Theory*.

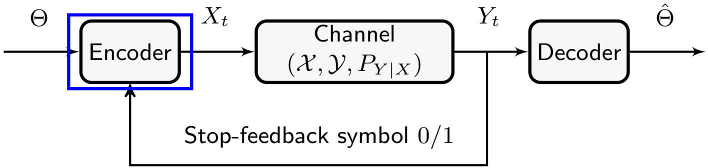
UCLA Samueli Electrical & Computer Engineering



Given $l > 0$, $n_1^m \in \mathbb{N}_+^m$ with $n_1 < n_2 < \cdots < n_m$, $M \in \mathbb{N}_+$, $\epsilon \in (0, 1)$, we want to specify an (l, n_1^m, M, ϵ) VLSF code.



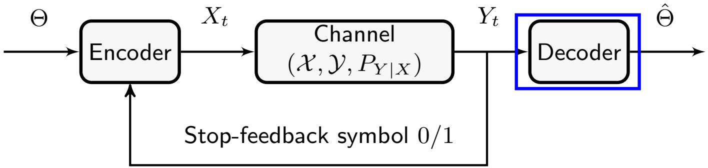
Codebook $U \in \mathcal{U}$: designed and fixed before transmission



Encoding function $e_t : \mathcal{U} \times [M] \to \mathcal{X}$:

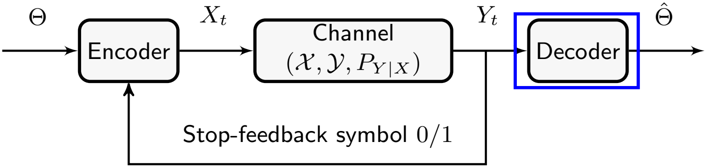
$$X_t = e_t(U, \Theta), \quad t \in \mathbb{N}_+$$

where $\Theta \sim \mathrm{Unif}\,([M])$.



Decoding function $g_t : \mathcal{U} \times \mathcal{Y}^t \to [M]$: providing the best estimate of Θ at time t, $t \in \{n_1, n_2, \ldots, n_m\}$.

UCLA Samueli Electrical & Computer Engineering



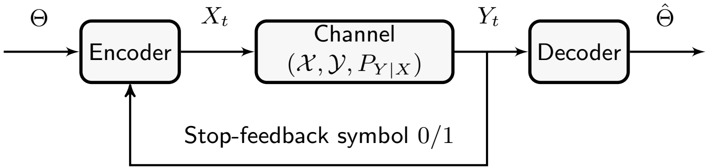
Stopping time $\tau \in \{n_i\}_{i=1}^m$: a function of filtration generated by $\{U, Y^{n_i}\}_{i=1}^m$ and must satisfy $\mathbb{E}[\tau] \leq l$.

Final decision: $\hat{\Theta} = g_\tau(Y^\tau)$

τ also needs to satisfy

$$P_e \triangleq \mathbb{P}[\Theta \neq \hat{\Theta}] \leq \epsilon.$$

UCLA Samueli Electrical & Computer Engineering



Goal: Determine $l^*(m, M, \epsilon) \triangleq \min\{l : \exists (l, n_1^m, M, \epsilon) \text{ VLSF code}\}$

Information density:

$$\iota(x^n; y^n) \triangleq \log \frac{\mathsf{P}_{Y^n|X^n}(y^n|x^n)}{\mathsf{P}_{Y^n}(y^n)} \tag{41}$$

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, Apr. 2010.

Information density:

$$\iota(x^n; y^n) \triangleq \log \frac{\mathsf{P}_{Y^n|X^n}(y^n|x^n)}{\mathsf{P}_{Y^n}(y^n)} \tag{41}$$

- If $\mathsf{P}_{X^n} = \prod_{i=1}^n \mathsf{P}_{X_i}$ and the channel is memoryless, $\iota(x^n; y^n) = \sum_{i=1}^n \iota(x_i; y_i)$.

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, Apr. 2010.

Information density:

$$\iota(x^n; y^n) \triangleq \log \frac{\mathsf{P}_{Y^n|X^n}(y^n|x^n)}{\mathsf{P}_{Y^n}(y^n)} \tag{41}$$

- If $\mathsf{P}_{X^n} = \prod_{i=1}^n \mathsf{P}_{X_i}$ and the channel is memoryless, $\iota(x^n; y^n) = \sum_{i=1}^n \iota(x_i; y_i)$.
- If $\mathsf{P}_X = \mathsf{P}_X^*$, define channel capacity and dispersion by

$$C \triangleq \mathbb{E}_{\mathsf{P}_X^* \mathsf{P}_{Y|X}}[\iota(X;Y)], \tag{42}$$

$$V \triangleq \mathbb{E}_{\mathsf{P}_X^* \mathsf{P}_{Y|X}}[\iota^2(X;Y)] - C^2. \tag{43}$$

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, Apr. 2010.

UCLA Samueli Electrical & Computer Engineering

Information density:

$$\imath(x^n;y^n) \triangleq \log \frac{\mathsf{P}_{Y^n|X^n}(y^n|x^n)}{\mathsf{P}_{Y^n}(y^n)} \tag{41}$$

- If $\mathsf{P}_{X^n} = \prod_{i=1}^n \mathsf{P}_{X_i}$ and the channel is memoryless, $\imath(x^n;y^n) = \sum_{i=1}^n \imath(x_i;y_i)$.
- If $\mathsf{P}_X = \mathsf{P}_X^*$, define channel capacity and dispersion by

$$C \triangleq \mathbb{E}_{\mathsf{P}_X^* \mathsf{P}_{Y|X}}[\imath(X;Y)], \tag{42}$$

$$V \triangleq \mathbb{E}_{\mathsf{P}_X^* \mathsf{P}_{Y|X}}[\imath^2(X;Y)] - C^2. \tag{43}$$

- We assume i.i.d. inputs $\sim P_X^*$.

Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, Apr. 2010.

UCLA Samueli Electrical & Computer Engineering

An Achievability Bound for (l, n_1^m, M, ϵ) VLSF Code

Theorem 14 (Yavas *et al.*, 2021)

Fix a constant $\gamma > 0$*, integer-valued decoding times* $n_1 < n_2 < \cdots < n_m$*, and a memoryless channel* $(\mathcal{X}, \mathcal{Y}, \mathsf{P}_{Y|X})$*. For any* $l > 0$ *and* $\epsilon \in (0,1)$*, there exists an* (l, n_1^m, M, ϵ') *VLSF code with*

$$l \leq n_m + \sum_{i=1}^{m-1} (n_i - n_{i+1}) \mathbb{P}\left[\bigcup_{j=1}^{i} \{\iota(X^{n_j}; Y^{n_j}) \geq \gamma\} \right], \tag{44}$$

$$\epsilon' \leq 1 - \mathbb{P}[\iota(X^{n_m}; Y^{n_m}) \geq \gamma] + (M-1)2^{-\gamma}, \tag{45}$$

where $\mathsf{P}_{X^{n_m}}$ *is the product of distributions of* m *subvectors of lengths* $n_i - n_{i-1}$*,* $i \in [m]$*, i.e.,*

$$\mathsf{P}_{X^{n_m}}(x_1^{n_m}) = \prod_{i=1}^{m} \mathsf{P}_{X_{n_{i-1}+1}^{n_i}}\left(x_{n_{i-1}+1}^{n_i}\right). \tag{46}$$

R. Yavas *et al.*, "Variable-length feedback codes with several decoding times for the Gaussian channel," *IEEE Int. Sym. Inf. Theory (ISIT)*, Jul. 2021.

UCLA Samueli Electrical & Computer Engineering

By relaxing $\mathbb{P}\left[\bigcup_{j=1}^{i}\{\iota(X^{n_j};Y^{n_j})\geq\gamma\}\right]$ to $\mathbb{P}[\iota(X^{n_i};Y^{n_i})\geq\gamma]$, define

$$N(\gamma,n_1^m)\triangleq n_m+\sum_{i=1}^{m-1}(n_i-n_{i+1})\mathbb{P}[\iota(X^{n_i};Y^{n_i})\geq\gamma], \tag{47}$$

$$\mathcal{F}_m(\gamma,M,\epsilon)\triangleq\{n_1^m\in\mathbb{R}_+^m: n_{i+1}-n_i\geq 1,\forall i\in[m-1]; \mathbb{P}[\iota(X^{n_m};Y^{n_m})\geq\gamma]\geq 1-\epsilon+(M-1)2^{-\gamma}\}. \tag{48}$$

By relaxing $\mathbb{P}\left[\bigcup_{j=1}^{i}\{\iota(X^{n_j};Y^{n_j})\geq\gamma\}\right]$ to $\mathbb{P}[\iota(X^{n_i};Y^{n_i})\geq\gamma]$, define

$$N(\gamma, n_1^m) \triangleq n_m + \sum_{i=1}^{m-1}(n_i - n_{i+1})\mathbb{P}[\iota(X^{n_i};Y^{n_i})\geq\gamma], \tag{47}$$

$$\begin{aligned}\mathcal{F}_m(\gamma, M, \epsilon) \triangleq \{n_1^m \in \mathbb{R}_+^m : n_{i+1} - n_i \geq 1, \forall i \in [m-1]; \\ \mathbb{P}[\iota(X^{n_m};Y^{n_m})\geq\gamma] \geq 1-\epsilon+(M-1)2^{-\gamma}\}.\end{aligned} \tag{48}$$

Integer program: for a given $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, $\epsilon \in (0,1)$, and $\gamma \geq \log\frac{M-1}{\epsilon}$,

$$\begin{aligned}\min_{n_1^m} \quad & N(\gamma, n_1^m) \\ \text{s.t.} \quad & n_1^m \in \mathcal{F}_m(\gamma, M, \epsilon) \\ & n_1^m \in \mathbb{N}_+^m.\end{aligned} \tag{49}$$

UCLA Samueli Electrical & Computer Engineering

By relaxing $\mathbb{P}\left[\bigcup_{j=1}^{i} \{\iota(X^{n_j};Y^{n_j}) \geq \gamma\}\right]$ to $\mathbb{P}[\iota(X^{n_i};Y^{n_i}) \geq \gamma]$, define

$$N(\gamma, n_1^m) \triangleq n_m + \sum_{i=1}^{m-1} (n_i - n_{i+1}) \mathbb{P}[\iota(X^{n_i};Y^{n_i}) \geq \gamma], \tag{47}$$

$$\begin{aligned}\mathcal{F}_m(\gamma, M, \epsilon) \triangleq \{n_1^m \in \mathbb{R}_+^m : n_{i+1} - n_i \geq 1, \forall i \in [m-1]; \\ \mathbb{P}[\iota(X^{n_m};Y^{n_m}) \geq \gamma] \geq 1 - \epsilon + (M-1)2^{-\gamma}\}.\end{aligned} \tag{48}$$

Integer program: for a given $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, $\epsilon \in (0,1)$, and $\gamma \geq \log \frac{M-1}{\epsilon}$,

$$\begin{aligned} \min_{n_1^m} \quad & N(\gamma, n_1^m) \\ \text{s.t.} \quad & n_1^m \in \mathcal{F}_m(\gamma, M, \epsilon) \\ & n_1^m \in \mathbb{N}_+^m. \end{aligned} \tag{49}$$

Two-step minimization: $\min_\gamma \min_{n_1^m} N(\gamma, n_1^m)$

UCLA Samueli Electrical & Computer Engineering

UCLA Samueli Electrical & Computer Engineering

Gaussian Model

For BI-AWGN channel,

- $\iota(x;Y) = 1 - \log\left(1 + e^{-2xY}\right)$ is continuous.
- $\iota(X^n;Y^n)$ is a sum of i.i.d. $\iota(X;Y)$.

UCLA Samueli Electrical & Computer Engineering

For BI-AWGN channel,

- $\iota(x;Y) = 1 - \log\left(1 + e^{-2xY}\right)$ is continuous.
- $\iota(X^n;Y^n)$ is a sum of i.i.d. $\iota(X;Y)$.

Central limit theorem (CLT): Let $W_1, W_2, \ldots, W_n$ be i.i.d. r.v.'s with zero mean, variance σ^2. Define the standardized sum

$$S_n \triangleq \frac{\sum_{i=1}^{n} W_i}{\sigma\sqrt{n}}. \tag{50}$$

Then, $\lim_{n\to\infty} \mathbb{P}[S_n \leq x] = \Phi(x)$.

For BI-AWGN channel,

- $\iota(x;Y) = 1 - \log\left(1 + e^{-2xY}\right)$ is continuous.
- $\iota(X^n;Y^n)$ is a sum of i.i.d. $\iota(X;Y)$.

Central limit theorem (CLT): Let $W_1, W_2, \ldots, W_n$ be i.i.d. r.v.'s with zero mean, variance σ^2. Define the standardized sum

$$S_n \triangleq \frac{\sum_{i=1}^{n} W_i}{\sigma\sqrt{n}}. \tag{50}$$

Then, $\lim_{n\to\infty} \mathbb{P}[S_n \leq x] = \Phi(x)$.

Gaussian Model [Wang *et al.*, 2017]: For n sufficiently large,

$$\mathbb{P}[\iota(X^n;Y^n) \geq \gamma] \approx Q\left(\frac{\gamma - nC}{\sqrt{nV}}\right). \tag{51}$$

H. Wang *et al.*, "An information density approach to analyzing and optimizing incremental redundancy with feedback", *IEEE Int. Sym. Inf. Theory (ISIT)*, Jun. 2017.

For BI-AWGN channel,

- $\iota(x;Y) = 1 - \log\left(1 + e^{-2xY}\right)$ is continuous.
- $\iota(X^n;Y^n)$ is a sum of i.i.d. $\iota(X;Y)$.

Central limit theorem (CLT): Let $W_1, W_2, \ldots, W_n$ be i.i.d. r.v.'s with zero mean, variance σ^2. Define the standardized sum

$$S_n \triangleq \frac{\sum_{i=1}^{n} W_i}{\sigma\sqrt{n}}. \tag{50}$$

Then, $\lim_{n\to\infty} \mathbb{P}[S_n \leq x] = \Phi(x)$.

Gaussian Model [Wang *et al.*, 2017]: For n sufficiently large,

$$\mathbb{P}[\iota(X^n;Y^n) \geq \gamma] \approx Q\left(\frac{\gamma - nC}{\sqrt{nV}}\right). \tag{51}$$

Question: What if n is small?

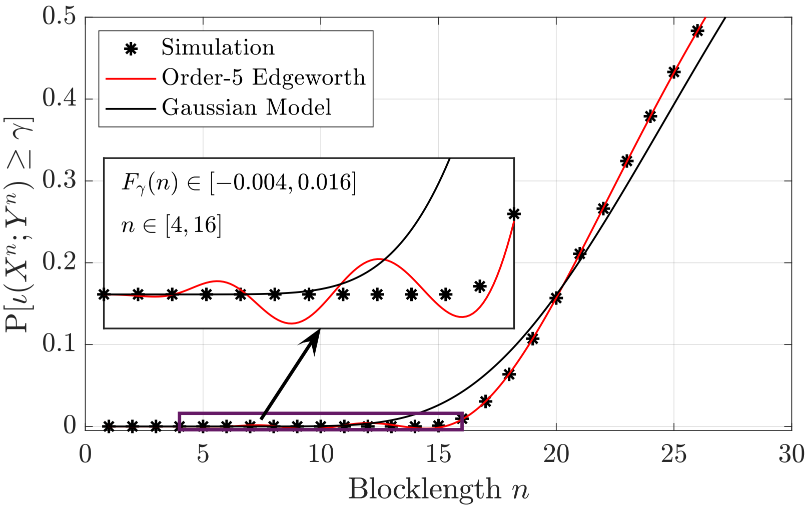
H. Wang *et al.*, "An information density approach to analyzing and optimizing incremental redundancy with feedback", *IEEE Int. Sym. Inf. Theory (ISIT)*, Jun. 2017.

Edgeworth expansion: Let $W_1, W_2, \ldots, W_n$ be i.i.d. absolutely continuous r.v.'s with zero mean, variance σ^2. Let $\{\kappa_i\}_{i=1}^{\infty}$ be the cumulants of W. If $\mathbb{E}[|W|^{s+2}] < \infty$ for some $s \in \mathbb{N}_+$, then,

$$\mathbb{P}[S_n \leq x] = \Phi(x) + \phi(x) \sum_{j=1}^{s} n^{-\frac{j}{2}} p_j(x) + o\left(n^{-\frac{s}{2}}\right), \tag{52}$$

where $p_j(x)$ requires cumulants $\kappa_3, \kappa_4, \ldots, \kappa_{j+2}$ of W.

F. Edgeworth, "The law of error," *Cambridge Philos. Trans.*, 1905.



Parameters setup: BI-AWGN channel at 0.2 dB, $\gamma = 13.62$.

UCLA Samueli Electrical & Computer Engineering

Petrov expansion: Let $W_1, W_2, \ldots, W_n$ be i.i.d. r.v.'s with zero mean, variance σ^2. Let $\{\kappa_i\}_{i=1}^{\infty}$ be the cumulants of W. If $x \geq 0$, $x = o(\sqrt{n})$, and $\mathbb{E}[e^{tW}] < \infty$ for $|t| < H$ for some $H > 0$,

$$\mathbb{P}[S_n \leq x] = 1 - Q(x) \exp\left\{\frac{x^3}{\sqrt{n}}\Lambda\left(\frac{x}{\sqrt{n}}\right)\right\}\left[1 + O\left(\frac{x+1}{\sqrt{n}}\right)\right], \tag{53}$$

$$\mathbb{P}[S_n \leq -x] = Q(x) \exp\left\{\frac{-x^3}{\sqrt{n}}\Lambda\left(\frac{-x}{\sqrt{n}}\right)\right\}\left[1 + O\left(\frac{x+1}{\sqrt{n}}\right)\right], \tag{54}$$

where $\Lambda(t) = \sum_{k=0}^{\infty} a_k t^k$ is called the Cramér series. Petrov provided

$$\Lambda^{[3]}(t) = \frac{\kappa_3}{6\kappa_2^{3/2}} + \frac{\kappa_4\kappa_2 - 3\kappa_3^2}{24\kappa_2^3}t + \frac{\kappa_5\kappa_2^2 - 10\kappa_4\kappa_3\kappa_2 + 15\kappa_3^3}{120\kappa_2^{9/2}}t^2 \tag{55}$$

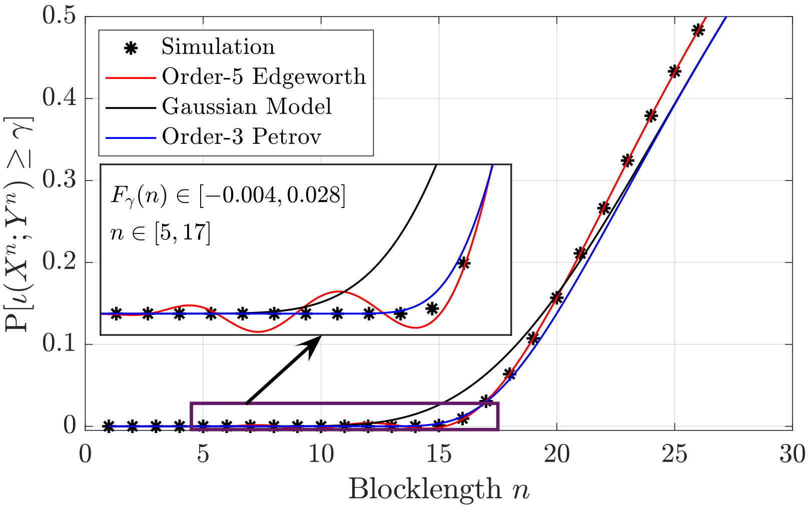
V. V. Petrov, "Sum of independent random variables," USA: Springer, Berlin, Heidelberg, 1975.

For BI-AWGN channel: $F_\gamma(n)$: a function to approximate $\mathbb{P}[\iota(X^n;Y^n) \geq \gamma]$.

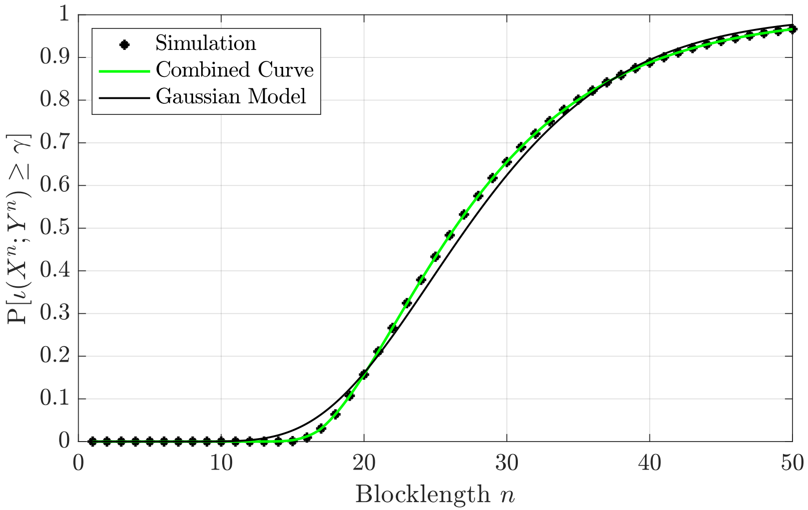
$$F_\gamma(n) = \begin{cases} Q\left(x(n)\right) - \phi(x(n)) \sum_{j=1}^{5} n^{-\frac{j}{2}} p_j(x(n)), \ n > n^* \\ Q(x(n)) \exp\left\{ \frac{x^3(n)}{\sqrt{n}} \Lambda^{[3]} \left(\frac{x(n)}{\sqrt{n}} \right) \right\}, \ 0 \leq n \leq n^*, \end{cases} \tag{56}$$

where

$$x(n) \triangleq \frac{\gamma - nC}{\sqrt{nV}}$$



Parameters setup: BI-AWGN channel at 0.2 dB, $\gamma = 13.62$, $n^* = 16.84$.



Parameters setup: BI-AWGN channel at 0.2 dB, $\gamma = 13.62$, $n^* = 16.84$.

Relaxed program: for a given $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, $\epsilon \in (0,1)$, and $\gamma \geq \log \frac{M-1}{\epsilon}$,

$$\begin{aligned} \min_{n_1^m} \quad & N(\gamma, n_1^m) \\ \text{s.t.} \quad & n_1^m \in \mathcal{F}_m(\gamma, M, \epsilon) \end{aligned} \tag{57}$$

UCLA Samueli Electrical & Computer Engineering

Theorem 15 (Gap-constrained SDO procedure)

Fix a memoryless channel $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *for which* $\iota(X;Y)$ *is continuous and* $\mathbb{P}[\iota(X^n;Y^n) \geq \gamma]$ *is increasing and differentiable. For a given* $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, $\epsilon \in (0,1)$, *and* $\gamma \geq \log \frac{M-1}{\epsilon}$, *the optimal real-valued decoding times* $n_1^*, n_2^*, \ldots, n_m^*$ *for the relaxed program* (57) *satisfy*

$$n_m^* = F_\gamma^{-1}\left(1 - \epsilon + (M-1)2^{-\gamma}\right), \quad (58)$$

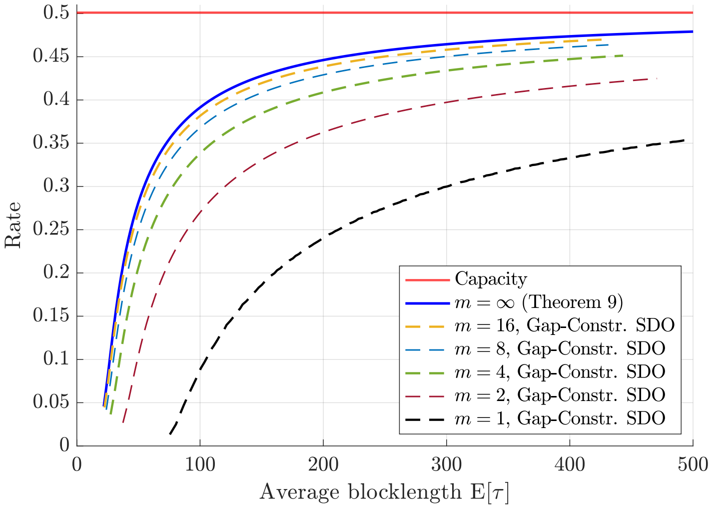
$$n_{i+1}^* = n_i^* + \max\left\{1, \frac{F_\gamma(n_i^*) - F_\gamma(n_{i-1}^*) - \lambda_{i-1}}{f_\gamma(n_i^*)}\right\}, \quad (59)$$

$$\lambda_i = \max\{\lambda_{i-1} + f_\gamma(n_i^*) - F_\gamma(n_i^*) + F_\gamma(n_{i-1}^*), 0\}, \quad (60)$$

where $i \in [m-1]$, $\lambda_0 \triangleq 0$, *and* $n_0^* \triangleq 0$.

UCLA Samueli Electrical & Computer Engineering

Achievability Bounds for (l, n_1^m, M, ϵ) VLSF Codes over BI-AWGN Channel



Parameters setup: BI-AWGN channel at 0.2 dB with $C = 0.5$. $\epsilon = 10^{-3}$.

What About BSC?

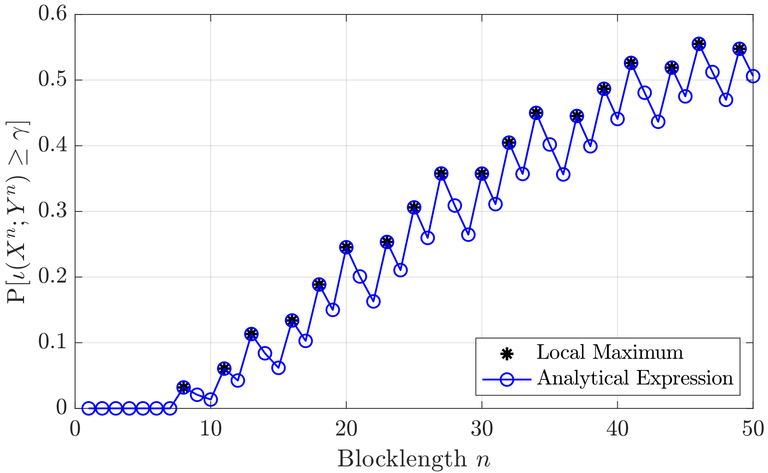
For BSC(p), $p \in (0, 1/2)$,

- $\iota(X;Y) = \log(2-2p) - Z\left(\log \frac{1-p}{p}\right)$ is a lattice r.v.

For BSC(p), $p \in (0, 1/2)$,

- $\iota(X;Y) = \log(2-2p) - Z\left(\log \frac{1-p}{p}\right)$ is a lattice r.v.
- The tail probability

$$\mathbb{P}[\iota(X^n;Y^n) \geq \gamma] = \mathbb{P}\left[\sum_{i=1}^{n} Z_i \leq \frac{n\log(2-2p)-\gamma}{\log\left((1-p)/p\right)}\right] = \sum_{c=0}^{\left\lfloor \frac{n\log(2-2p)-\gamma}{\log((1-p)/p)} \right\rfloor} \binom{n}{c} p^c (1-p)^{n-c}. \tag{61}$$



Parameters setup: BSC(0.35), $\gamma = 3$.

UCLA Samueli Electrical & Computer Engineering

Theorem 16

Fix $\gamma > 0$ *and* $p \in (0, 1/2)$. *Define* $\alpha_i \triangleq \left\lceil \frac{\gamma + i \log((1-p)/p)}{\log(2-2p)} \right\rceil$, $i \in \mathbb{N}$. *Then,*

$$\mathbb{P}[\imath(X^n; Y^n) \geq \gamma] < \mathbb{P}[\imath(X^{n+1}; Y^{n+1}) \geq \gamma], \quad \text{if } n = \alpha_i - 1, \tag{62}$$

$$\mathbb{P}[\imath(X^n; Y^n) \geq \gamma] > \mathbb{P}[\imath(X^{n+1}; Y^{n+1}) \geq \gamma], \quad \text{if } n \in [\alpha_i, \alpha_{i+1} - 1). \tag{63}$$

α_i *is called the* i*th local maximizer*, $i \in \mathbb{N}$.

Theorem 16

Fix $\gamma > 0$ *and* $p \in (0, 1/2)$. *Define* $\alpha_i \triangleq \left\lceil \frac{\gamma + i \log((1-p)/p)}{\log(2-2p)} \right\rceil$, $i \in \mathbb{N}$. *Then,*

$$\mathbb{P}[\imath(X^n; Y^n) \geq \gamma] < \mathbb{P}[\imath(X^{n+1}; Y^{n+1}) \geq \gamma], \quad \text{if } n = \alpha_i - 1, \tag{62}$$

$$\mathbb{P}[\imath(X^n; Y^n) \geq \gamma] > \mathbb{P}[\imath(X^{n+1}; Y^{n+1}) \geq \gamma], \quad \text{if } n \in [\alpha_i, \alpha_{i+1} - 1). \tag{63}$$

α_i *is called the* i*th* *local maximizer*, $i \in \mathbb{N}$.

Lemma 3

Fix a BSC(p), $p \in (0, 1/2)$. *For a given* $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, *and* γ, *if* $m < n_m^*(M, \gamma)$, *the optimal decoding times* n_1^m *for minimizing* $N(\gamma, n_1^m)$ *are among* $\{\alpha_i\}_{i=1}^{\infty}$.

UCLA Samueli Electrical & Computer Engineering

Define

$$n_m^* \triangleq \min\{n \in \mathbb{N} : \mathbb{P}[S_{n_m} \geq \gamma] \geq 1 - \epsilon + (M-1)2^{-\gamma}\}. \tag{64}$$

$$g_-^{(1)}(n_1) \triangleq \max_{\substack{n \in [1, n_m^* - m + 1] \\ \mathbb{P}[S_n \geq \gamma] < \mathbb{P}[S_{n_1} \geq \gamma]}} \frac{\mathbb{P}[S_n \geq \gamma](n_1 - n)}{\mathbb{P}[S_{n_1} \geq \gamma] - \mathbb{P}[S_n \geq \gamma]}, \tag{65}$$

$$g_+^{(1)}(n_1) \triangleq \min_{\substack{n \in [1, n_m^* - m + 1] \\ \mathbb{P}[S_n \geq \gamma] > \mathbb{P}[S_{n_1} \geq \gamma]}} \frac{\mathbb{P}[S_n \geq \gamma](n_1 - n)}{\mathbb{P}[S_{n_1} \geq \gamma] - \mathbb{P}[S_n \geq \gamma]}. \tag{66}$$

$$g_-^{(i)}(n_i, n_{i-1}) \triangleq \max_{\substack{n \in [n_{i-1}+1, n_m^* - m + i] \\ \mathbb{P}[S_n \geq \gamma] < \mathbb{P}[S_{n_i} \geq \gamma]}} \frac{\mathbb{P}[S_n \geq \gamma] - \mathbb{P}[S_{n_{i-1}} \geq \gamma]}{\mathbb{P}[S_{n_i} \geq \gamma] - \mathbb{P}[S_n \geq \gamma]}(n_i - n), \tag{67}$$

$$g_+^{(i)}(n_i, n_{i-1}) \triangleq \min_{\substack{n \in [n_{i-1}+1, n_m^* - m + i] \\ \mathbb{P}[S_n \geq \gamma] > \mathbb{P}[S_{n_i} \geq \gamma]}} \frac{\mathbb{P}[S_n \geq \gamma] - \mathbb{P}[S_{n_{i-1}} \geq \gamma]}{\mathbb{P}[S_{n_i} \geq \gamma] - \mathbb{P}[S_n \geq \gamma]}(n_i - n). \tag{68}$$

- $g_-^{(i)}(\cdot) = -\infty$ if the maximizer is empty.
- $g_+^{(i)}(\cdot) = \infty$ if the minimizer is empty

UCLA Samueli Electrical & Computer Engineering

Theorem 17 (Discrete SDO procedure)

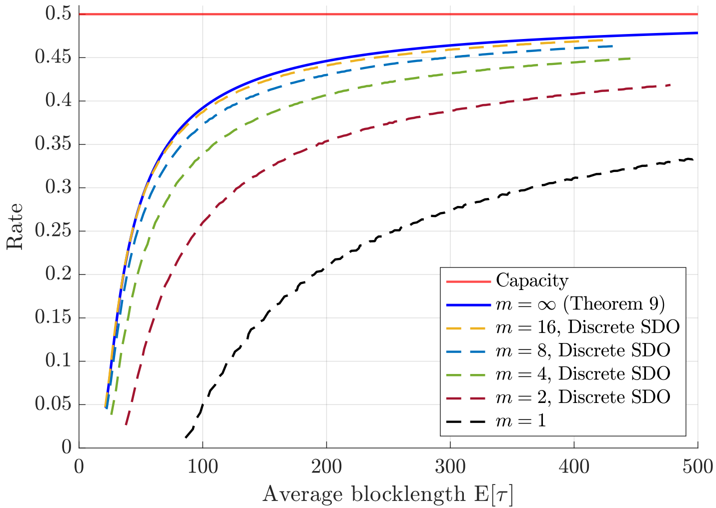
Fix a memoryless channel $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ *and scalars* $m \in \mathbb{N}_+$, $M \in \mathbb{N}_+$, $\epsilon \in (0,1)$, *and* $\gamma \geq \log \frac{M-1}{\epsilon}$. *Define* $S_n \triangleq \iota(X^n; Y^n)$. *The optimal integer-valued decoding times* $n_1^*, n_2^*, \dots, n_m^*$ *for the integer program* (49) *satisfy*

$$n_1^* + \max(1, g_-^{(1)}(n_1^*)) \leq n_2^* \leq n_1^* + g_+^{(1)}(n_1^*), \tag{69}$$

$$n_2^* + \max(1, g_-^{(2)}(n_2^*, n_1^*)) \leq n_3^* \leq n_2^* + g_+^{(2)}(n_2^*, n_1^*), \tag{70}$$

$$\dots$$

$$n_{m-1}^* + \max(1, g_-^{(m-1)}(n_{m-1}^*, n_{m-2}^*)) \leq n_m^* \leq n_{m-1}^* + g_+^{(m-1)}(n_{m-1}^*, n_{m-2}^*), \tag{71}$$



Parameters setup: BSC(0.11) with $C = 0.5$. $\epsilon = 10^{-3}$.

UCLA Samueli Electrical & Computer Engineering

Case 1: BI-AWGN channel with no feedback

- We designed a new block code called CRC-aided convolutional code.
- Simulation shows that several CRC-TBCCs under SLVD outperforms the RCU bound at a reasonable complexity.

Case 1: BI-AWGN channel with no feedback

- We designed a new block code called CRC-aided convolutional code.
- Simulation shows that several CRC-TBCCs under SLVD outperforms the RCU bound at a reasonable complexity.

Case 2: Binary channels with full, noiseless feedback

- For BSC, we developed refined non-asymptotic VLF achievability bound that outperforms Polyanskiy's VLSF achievability bound.
- For BAC, we generalized the SED coding scheme and developed a non-asymptotic VLF achievability bound.

UCLA Samueli Electrical & Computer Engineering

Case 1: BI-AWGN channel with no feedback

- We designed a new block code called CRC-aided convolutional code.
- Simulation shows that several CRC-TBCCs under SLVD outperforms the RCU bound at a reasonable complexity.

Case 2: Binary channels with full, noiseless feedback

- For BSC, we developed refined non-asymptotic VLF achievability bound that outperforms Polyanskiy's VLSF achievability bound.
- For BAC, we generalized the SED coding scheme and developed a non-asymptotic VLF achievability bound.

Case 3: Binary-input channels with finite, stop feedback

- We developed two methods to evaluate the VLSF achievability bounds.
- For both BI-AWGN channel and BSC, Polyanskiy's VLSF achievability bound can be approached with a small number of decoding times.

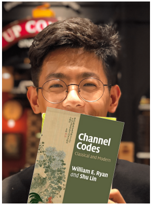
UCLA Samueli Electrical & Computer Engineering

Acknowledgment

I would like to thank my significant collaborators:

Ethan Liang

Minghao Pan



Linfang Wang

Recep Yavas

UCLA Samueli Electrical & Computer Engineering

I would also like to thank

- my advisor: Prof. Richard Wesel;
- my committee members: Profs. Lara Dolecek, Christina Fragouli, Alexander Sherstov, and Dariush Divsalar;
- all faculty members of the ECE Department and Mathematics Department;
- my friends in California: Weinan Song, Linfang Wang, Bijie Bai, Yun Liao, Minghao Pan, and many others;
- National Science Foundation (NSF) and Qualcomm for funding support.

UCLA Samueli Electrical & Computer Engineering

Thank you!

UCLA Samueli Electrical & Computer Engineering