Non-Asymptotic Analysis of Single-Receiver Channels with Limited Feedback

Recep Can Yavas
Advisors: Prof. Michelle Effros and Prof. Victoria Kostina
California Institute of Technology

Ph.D. Defense
Aug. 29, 2022

Outline

Outline

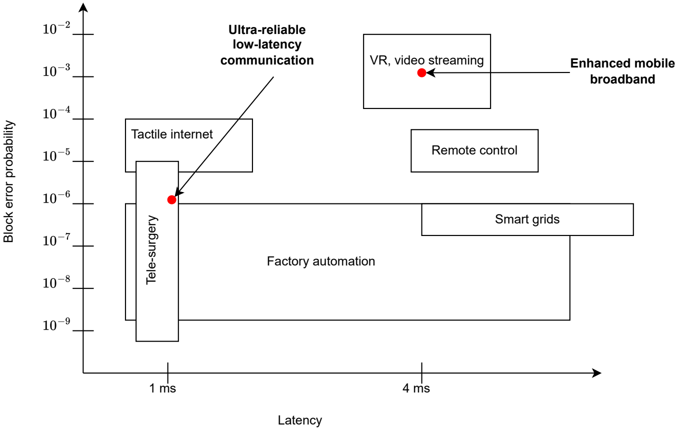
1. Motivation
2. Dissertation
   - Moderate Deviations Analysis of Point-to-Point Channels
   - Variable-Length Sparse Stop-Feedback Codes
   - Random Access Channels with Sparse Stop-Feedback
3. Conclusion and Future Directions

5G and newer generation networks offer three new service categories

- Ultra-reliable low-latency communication
  - Delay-sensitive, mission critical applications (tele-surgery, factory automation)
- Enhanced mobile broadband
  - High bandwidth and low power consumption requirements (HD video streaming, virtual reality)
- Massive machine-type communication
  - Massive and unknown number of communicating devices with low power consumption (Internet of Things, smart cities, smart grids)

Goal: Design coding schemes that are suitable for applications like these, analyze their performance, and understand the fundamental limits from an information-theoretic perspective.

Applications with Stringent Requirements



1 ms latency $\approx n = 100$–200

M. Shirvanimoghaddam et al., "Short Block-Length Codes for Ultra-Reliable Low Latency Communications", IEEE Commun. Mag., Feb. 2019.

Outline

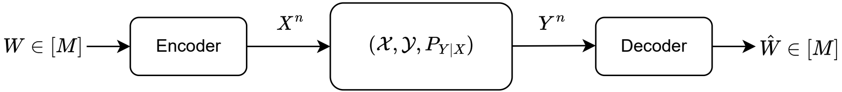
1. Motivation

2. Dissertation
   - Moderate Deviations Analysis of Point-to-Point Channels
   - Variable-Length Sparse Stop-Feedback Codes
   - Random Access Channels with Sparse Stop-Feedback

3. Conclusion and Future Directions

Outline

Channel Coding Problem



- M = codebook size, n = blocklength, ϵ = average error probability, $[M] \triangleq \{1, \ldots, M\}$.
- Fundamental limit: $M^*(n, \epsilon)$ = maximum achievable codebook size compatible with n and ϵ
- Goal: Compute $M^*(n, \epsilon)$ and find the optimal codes that achieve it
- Problem: Exact computation of $M^*(n, \epsilon)$ is intractable
- Solution: Derive non-asymptotic bounds

$$\underline{M}^*(n, \epsilon) \leq M^*(n, \epsilon) \leq \overline{M}^*(n, \epsilon)$$

- Example of non-asymptotic bound [Polyanskiy 10', random coding union bound]:

$$\underline{M}^*(n,\epsilon) = \max\{M : \mathbb{E}\left[\min\{1, (M-1)\mathbb{P}\left[\imath(\bar{X}^n; Y^n) \geq \imath(X^n; Y^n) | X^n, Y^n\right]\}\right] \leq \epsilon\}$$

  The "solution" is an optimization problem
  - it does not make clear how $M^*(n,\epsilon)$ varies with $(P_{Y|X}, n, \epsilon)$
  - even approximately computing it is nontrivial

- We are interested in asymptotic expansions (Taylor-like expansions of $\frac{\log M^*(n,\epsilon)}{n}$ around capacity) that
  - are tight for $(P_{Y|X}, n, \epsilon)$ of interest
  - inform about how $M^*(n,\epsilon)$ varies with $(P_{Y|X}, n, \epsilon)$
  - cover a variety of channels

- Analyzing $M^*(n,\epsilon)$ as $n \to \infty$ is useful!

Error Probability Regimes in Channel Coding

As $n \to \infty$, we need to specify how ϵ is related to n

- **Central limit theorem (CLT) regime:**

$$\begin{aligned} &\epsilon \in (0,1) \\ &\log M^*(n,\epsilon) = nC - O(\sqrt{n}) \\ &C = \text{capacity} \end{aligned}$$

- **Moderate deviations (MD) regime**

$$\epsilon \to 0 \text{ and } -\frac{1}{n}\log\epsilon \to 0 \qquad \log M^*(n,\epsilon) = nC - g(n), \quad \sqrt{n} \ll g(n) \ll n$$

- **Large deviations (LD) regime:**

$$\begin{aligned} &\epsilon \approx e^{-nE(R)} \\ &\log M^*(n,\epsilon) = nR, \quad R < C \end{aligned}$$

MD regime is relevant to ultra-reliable short-blocklength regime where both ϵ and n are small.

CLT and LD regimes fail to capture this regime.

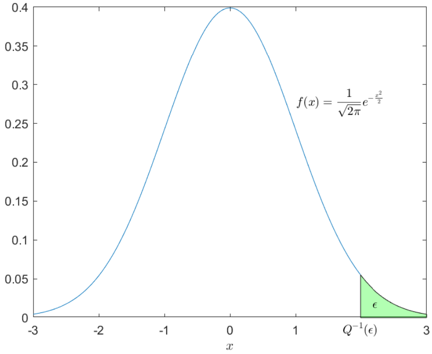
- [Strassen 62'] and [Polyanskiy et al. 10'] show

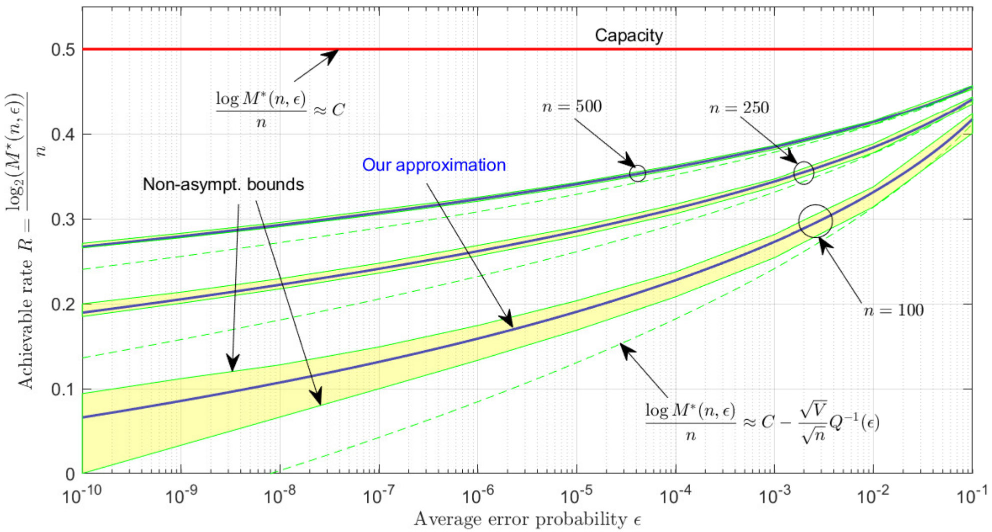
$$\log M^*(n,\epsilon) = \underbrace{nC - \sqrt{nV}Q^{-1}(\epsilon)}_{\text{Gaussian approximation}} + O(\log n)$$

C = capacity

V = dispersion (a quantity that depends only on $P_{Y|X}$)

$Q^{-1}(\cdot)$ = inverse of $Q(\cdot)$, complementary cumulative distribution function of $\mathcal{N}(0,1)$





Achievable rates for BSC(0.11)

Non-Gaussianity:

$$\zeta(n,\epsilon) \triangleq \log M^*(n,\epsilon) - \underbrace{\left(nC - \sqrt{nV}Q^{-1}(\epsilon)\right)}_{\text{Gaussian approximation}}$$

1. **CLT**: [Polyanskiy et al. 10'] (achievability) and [Tomamichel-Tan 13'] (converse): for $\underbrace{\text{nonsingular channels}}_{\text{excludes a family of erasure channels}}$

$$\zeta(n,\epsilon) = \frac{1}{2}\log n + O(1)$$

2. **CLT**: [Moulin 17']: for nonsingular channels with a regularity condition (excludes the BSC)

$$\zeta(n,\epsilon) \geq \frac{1}{2}\log n + \underline{S}Q^{-1}(\epsilon)^2 + \underline{B} + o(1)$$
$$\zeta(n,\epsilon) \leq \frac{1}{2}\log n + \overline{S}Q^{-1}(\epsilon)^2 + \overline{B} + o(1)$$

where $\underline{S}, \overline{S}, \underline{B}$, and $\overline{B}$ depend only on $P_{Y|X}$

- **MD**: [Altuğ and Wagner 14']:

$$\zeta(n, \epsilon_n) = o(\sqrt{n} Q^{-1}(\epsilon_n))$$

Gaussian approximation (without $O(\log n)$ term) is valid in the MD regime

- **Question:** How does the non-Gaussianity $\zeta(n, \epsilon_n)$ scale in the MD regime?

Channel Skewness

- A new operational quantity: Channel skewness

$$S \triangleq \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{\zeta(n,\epsilon) - \frac{1}{2}\log n}{(Q^{-1}(\epsilon))^2}$$

We are expecting:

$$\log M^*(n,\epsilon_n) = nC - \sqrt{nV}Q^{-1}(\epsilon_n) + \frac{1}{2}\log n + SQ^{-1}(\epsilon_n)^2 + o(Q^{-1}(\epsilon_n)^2)$$

- Information density: $\imath(X;Y) \triangleq \log \frac{P_{Y|X}(Y|X)}{P_Y(Y)}$ under some fixed input distribution P_X

$$\left.\begin{array}{ll} C & = \mathbb{E}\left[\imath(X;Y)\right] \\ V & = \mathrm{Var}\left[\imath(X;Y)\right] \end{array}\right\} \text{ under capacity-achieving } P_X$$

$$S = \text{governed by the third central moment of information density}$$

Our Bounds on the non-Gaussianity in the MD Regime

ϵ_n in MD regime: $-\frac{1}{n}\log\epsilon_n \to 0$, $\epsilon_n \to 0$

Theorem 1

MD: *for nonsingular discrete memoryless point-to-point channels*

$$\zeta(n,\epsilon_n) \geq \frac{1}{2}\log n + \underline{S}\, Q^{-1}(\epsilon_n)^2 + O\left(\frac{Q^{-1}(\epsilon_n)^3}{\sqrt{n}}\right) + O(1)$$

$$\zeta(n,\epsilon_n) \leq \frac{1}{2}\log n + \overline{S}\, Q^{-1}(\epsilon_n)^2 + O\left(\frac{Q^{-1}(\epsilon_n)^3}{\sqrt{n}}\right) + O(1)$$

- $\underline{S}$ and $\overline{S}$ are the same as Moulin's bounds, and our results apply to a wider class of channels
- $\underline{S}$ and $\overline{S}$ are easy-to-compute constants that depend only on the channel $P_{Y|X}$ and are related to $\imath(X;Y)$
- $\underline{S}$ and $\overline{S}$ give lower and upper bounds for channel skewness S
- $\frac{1}{2}\log n + O(1)$ is no longer accurate because as $n\to\infty$, $\epsilon_n \to 0$ and $Q^{-1}(\epsilon_n)^2 \to \infty$

- Lower bound is achieved using i.i.d. random codewords drawn from a distribution P_X that is shifted from the capacity-achieving P_X^*

$$P_X = P_X^* + \frac{Q^{-1}(\epsilon_n)}{\sqrt{n}}\mathbf{h}$$

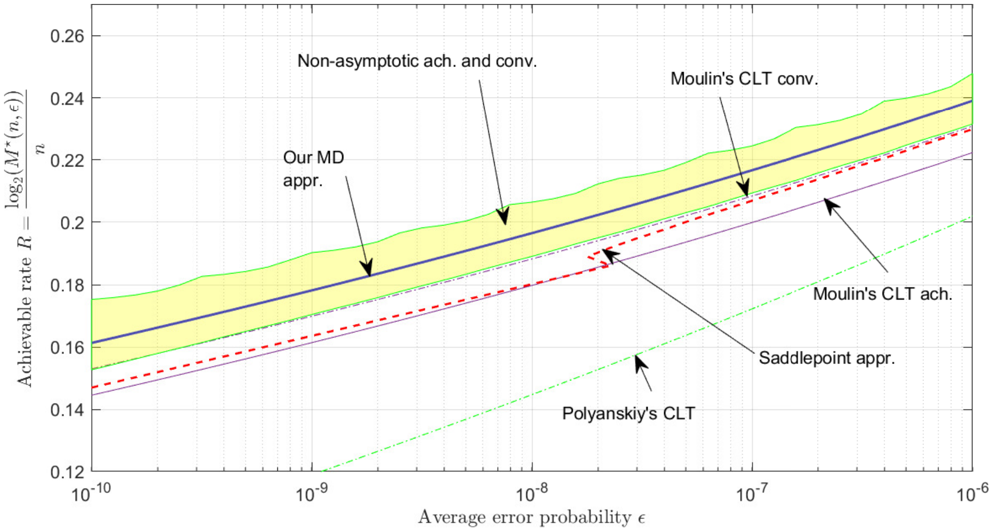
$\mathbf{h}$ = a vector that is a function of $P_{Y|X}$

- For Cover-Thomas symmetric channels (all rows (resp. columns) are permutation of each other), e.g., BSC

$$S = \overline{S} = \underline{S} = \frac{\mathrm{Sk}(P_X^*)\sqrt{V}}{6} + \frac{1}{2}$$
$$P_X = P_X^*$$
$$\mathrm{Sk}(P_X^*) = \text{skewness of } \imath(X;Y) \text{ at } P_X^*$$

- For the Gaussian channel, we compute the channel skewness $S(P)$ exactly as a function of power P

Comparison of Our Approximation to Existing Work



Achievable rates for BSC(0.11), $n = 200$.

- We relax the random coding union bound from [Polyanskiy et al. 10'] as

$$P_{\text{error}} \leq \mathbb{E}\left[\min\{1, (M-1)\mathbb{P}\left[\imath(\bar{X}^n; Y^n) \geq \imath(X^n; Y^n) | X^n, Y^n\right]\}\right]$$
$$\leq \mathbb{P}\left[\imath(X^n; Y^n) < \tau\right] + (M-1)\mathbb{P}\left[\imath(\bar{X}^n; Y^n) \geq \imath(X^n; Y^n) \geq \tau\right]$$

where $\bar{X}^n$ is a sample from the codebook that is independent of Y^n.

  - Relaxation of the ML decoder (information density threshold rule + ML rule)
  - Advantageous because we can analyze each probability term separately

Achievability Proof

- The optimal allocation of ϵ_n (equivalently, the optimal τ) is:

$$\underbrace{\mathbb{P}\left[\imath(X^n;Y^n) < \tau\right]}_{\epsilon_n - \epsilon_n \frac{Q^{-1}(\epsilon_n)}{\sqrt{nV(P_X)}}} + \underbrace{(M-1)\mathbb{P}\left[\imath(\bar{X}^n;Y^n) \geq \imath(X^n;Y^n) \geq \tau\right]}_{\epsilon_n \frac{Q^{-1}(\epsilon_n)}{\sqrt{nV(P_X)}}}$$

- $\mathbb{P}\left[\imath(X^n;Y^n) < \tau\right]$: MD regime of probability theory
  $\mathbb{P}\left[\imath(\bar{X}^n;Y^n) \geq \imath(X^n;Y^n) \geq \tau\right]$: LD regime of probability theory

Moderate Deviations Theorem

Theorem 2 (Petrov Expansion)

Let $X_1, \ldots, X_n$ *be i.i.d.,* $\mathbb{E}[X_1] = 0$ *and* $\mathrm{Var}[X] = \sigma^2$. *Let* $x = o(\sqrt{n})$, *and* $x \geq 0$. *Under Cramér's condition, i.e.,* $\mathbb{E}\left[e^{tX}\right] < \infty$ *in the neighborhood of zero,*

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq \sqrt{n\sigma^2}x\right] = Q(x)\exp\left(\frac{x^3 \mathrm{Sk}(X_1)}{6\sqrt{n}} + O\left(\frac{x^4}{n}\right)\right)\left(1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right)$$

where $\mathrm{Sk}(X_1) \triangleq \frac{\mathbb{E}[X_1^3]}{\sigma^3}$.

- It is an asymptotic equality
- It has a multiplicative correction term to Gaussian approximation $Q(x)$
- Skewness $\mathrm{Sk}(X_1)$ dominates the correction term

- $\mathbb{P}\left[\imath(X^n;Y^n) < \tau\right]$: Apply Petrov expansion
  $\mathbb{P}\left[\imath(\bar{X}^n;Y^n) \geq \imath(X^n;Y^n) \geq \tau\right]$: Apply strong large deviations theorem to get

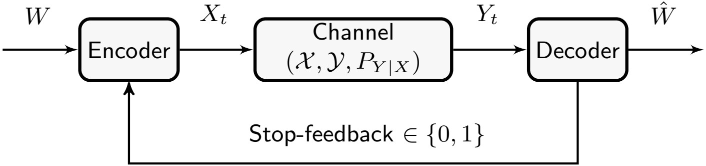
$$\log M^* \geq nI(P_X) - \sqrt{nV(P_X)}Q^{-1}(\epsilon_n) + \frac{1}{2}\log n$$
$$+ Q^{-1}(\epsilon_n)^2\left(\frac{\mathrm{Sk}(P_X)\sqrt{V(P_X)}}{6} + \frac{1-\eta(P_X)}{2(1+\eta(P_X))}\right) + O\left(\frac{Q^{-1}(\epsilon_n)^3}{\sqrt{n}}\right) + O(1)$$

- Take Taylor series expansions around the capacity-achieving P_X^*.

Summary of Part 1

1. In the MD regime of channel coding:
   - We derive easy-to-compute lower and upper bounds to the non-Gaussianity
   - For symmetric channels, the bounds match each other, and channel skewness S is governed by the skewness of information density $\imath(X;Y)$
   - Our MD approximation is more accurate than the CLT approximation of Polyanskiy, especially for lower error probabilities (e.g., $\epsilon \leq 10^{-6}$)
2. For some non-symmetric channels, capacity-achieving distribution does not achieve S.
3. We compute the skewness of the Gaussian channel.

Outline



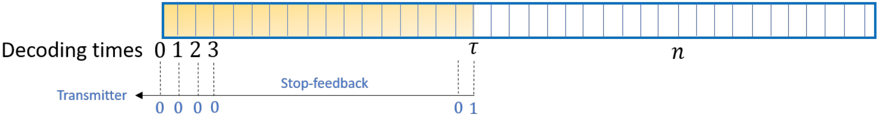
Variable-length: Decoding occurs at a random time depending on channel outputs

Stop-feedback: 1 bit feedback whenever a decoding attempt is made

Higher reliability: Earlier decoding when the noise is low, later decoding when the noise is high

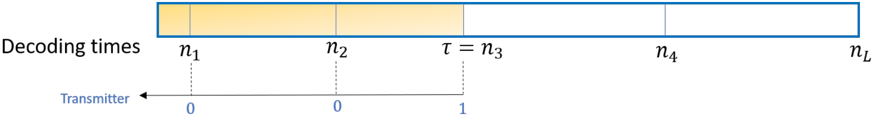
Why Sparse Stop-Feedback?

- VLSF code with $L = \infty$ decoding times [Polyanskiy et al. 11'], [Burnashev 76']: Impractical

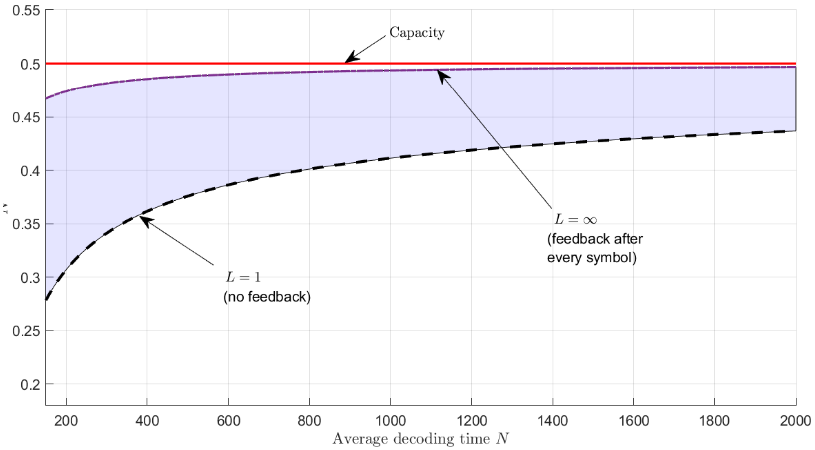


  - Transmitter constantly listens to the feedback signal $\Longrightarrow$ High power consumption
  - Half duplex devices cannot transmit and receive signals at the same time $\Longrightarrow$ Lowers achievable rates due to round trip delay
  - Practical codes such has HARQ schemes employ incremental redundancy which has sporadic feedback

- VLSF code with constant L decoding times (this work)

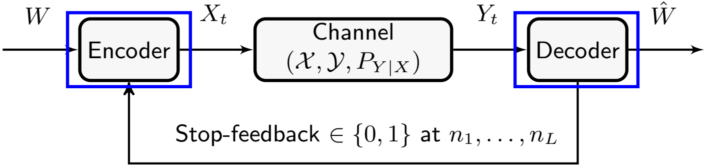


  - Feedback is sporadic



$L = \#$ of available decoding times

What rate can we achieve with several decoding times?

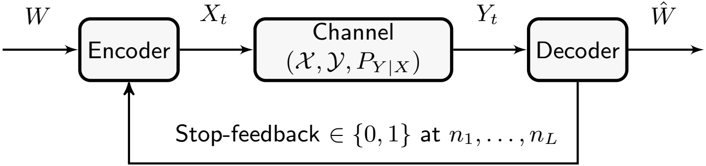


Encoding function $\mathsf{f}_t \colon [M] \to \mathcal{X}$:

$$X_t = \mathsf{f}_t(W), \quad t \in \mathbb{N}_+$$

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

Decoding function $\mathsf{g}_t \colon \mathcal{Y}^t \to [M]$: provides the estimate of W at time $t \in \{n_1, n_2, \ldots, n_L\}$.



Stopping time: $\tau \in \{n_i\}_{i=1}^{L}$ adapted to the filtration generated by $\{Y^{n_i}\}_{i=1}^{L}$

Decoded message: $\hat{W} = \mathrm{g}_\tau(Y^\tau)$

Goal: Find

$$\begin{aligned} M^*(N, L, \epsilon) \triangleq \max_{n_1, \ldots, n_L} \quad & M \\ \text{s.t.} \quad & \mathbb{E}[\tau] \leq N \\ & \mathbb{P}\left[\hat{W} \neq W\right] \leq \epsilon \end{aligned}$$

$L = \infty$:

- [Burnashev 76']: error exponent $\lim\limits_{\epsilon \to 0} \frac{-\log \epsilon}{\mathbb{E}[\tau]}$ for the discrete memoryless point-to-point channel as $N \to \infty$
- [Polyanskiy et al. 11']: VLSF codes in the CLT regime

$$\frac{NC}{1-\epsilon} - \log N + O(1) \leq \log M^*(N, \infty, \epsilon) \leq \frac{NC}{1-\epsilon} + O(1)$$

$$\log M^*(N, 1, \epsilon) = NC - \sqrt{NV} Q^{-1}(\epsilon) + O(\log N)$$

Constant L**:**

- [Vakilinia et al. 16']: VLSF codes with L decoding times over the binary-input Gaussian channel
- They estimate the statistics of τ through simulation.
- They do not solve the problem analytically $\Longrightarrow$ no second-order analysis

Theorem 3 (Achievability)

Fix $L \geq 2$, $\epsilon \in (0,1)$. For discrete memoryless point-to-point channels, it holds that

$$\log M^*(N, L, \epsilon) \geq \frac{NC}{1-\epsilon} - \sqrt{N \log_{(L-1)}(N) \frac{V}{1-\epsilon}} + o(\sqrt{N})$$

$$\log_{(L)}(\cdot) \triangleq \overbrace{\log(\log(\ldots(\log(\cdot))))}^{L \text{ times}}$$

- Proof analyzes a non-asymptotic bound
- By using refined probability tools and KKT conditions, we approximately optimize the non-asymptotic bound with respect to the choices of $n_1, \ldots, n_L$

Performance of VLSF Codes with L Decoding Times

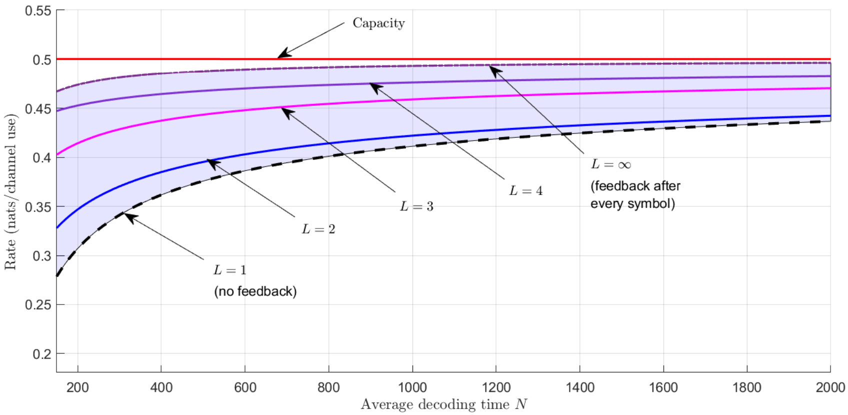


Figure 1: BSC(0.11), $\epsilon = 10^{-3}$

Diminishing performance improvement as L increases!

Sequential Hypothesis Tests (SHTs)

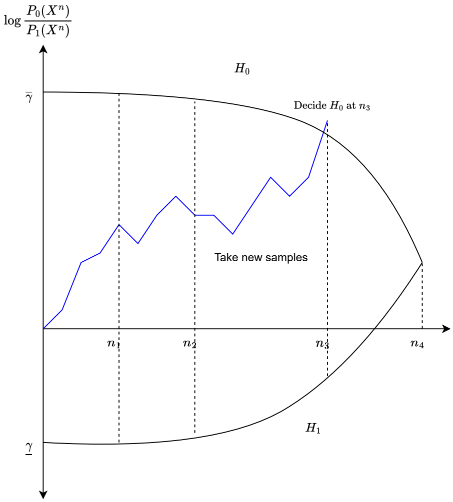
- Consider two hypotheses for the distribution of X^∞

$$H_0 : X^\infty \sim P_0^\infty$$
$$H_1 : X^\infty \sim P_1^\infty$$

- At each available decoding time, an SHT chooses between

$$\{\text{Decide } H_0, \text{Decide } H_1, \text{Take new sample(s)}\}$$

- Wald's theorem: the optimal SHT is a two-sided threshold test that uses log-likelihood ratio (sequential version of Neyman-Pearson lemma)



- Construct the SHT

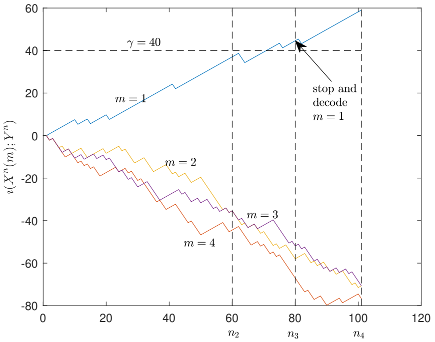
$H_0 : (X,Y) \sim P_X P_{Y|X} \Longrightarrow$ stop and decode

$H_1 : (X,Y) \sim P_X P_Y \Longrightarrow$ eliminate from decoding

where P_X is capacity-achieving and run it for all M messages separately

- At time $n_1 = 0$:
  - With probability $p = \epsilon - \frac{1}{\sqrt{N \log N}}$, declare H_1 for all $m \in [M]$
  - With probability $1 - p$, pass $n_1 = 0$ without decoding
- If $n_1 = 0$ is passed, then particularize the SHT to information-density threshold test:

$$\tau_m = \min\{n \in \{n_2, \dots, n_L\} : \imath(X^n(m); Y^n) \geq \gamma\}$$



- Analyzing this scheme gives us a non-asymptotic bound. Then, we optimize over $n_1, \dots, n_L$ using KKT conditions.

Maximal power constraint:

$$\|\mathsf{f}(m)^{n_\ell}\|_2^2 \leq n_\ell P \qquad \forall \ell \in [L], m \in [M]$$

Theorem 4

Fix $L \geq 2$, $P > 0$, *and* $\epsilon \in (0,1)$.

$$\log M^*_{\max}(N, L, \epsilon, P) \geq \frac{NC(P)}{1-\epsilon} - \sqrt{N \log_{(L-1)}(N) \frac{V(P)}{1-\epsilon}} + o(\sqrt{N})$$

where

$$C(P) = \frac{1}{2}\log(1+P), \quad V(P) = \frac{P(P+2)}{2(P+1)^2}$$

- **Encoder:** we generate codewords uniformly at random over a restricted subset on n_L-dimensional sphere

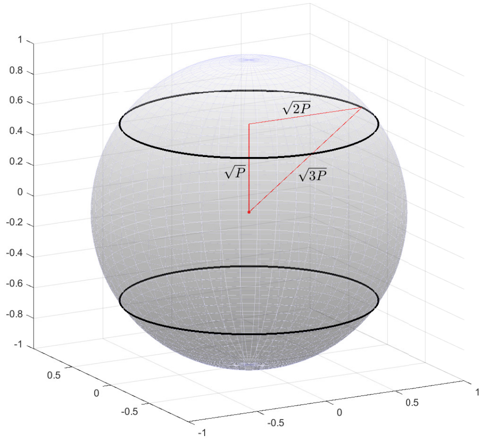
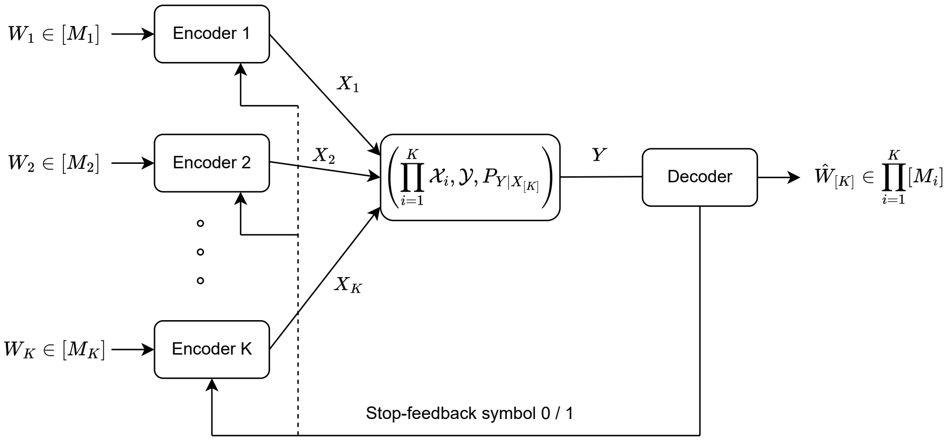


Figure 2: $L = 2$, $n_1 = 2, n_2 - n_1 = 1$



- $\{n_1, \ldots, n_L\}$ = the set of available decoding times
- $\tau \in \{n_1, \ldots, n_L\}$ = stopping time
- $X_{[K]} \triangleq X_1, \ldots, X_K$

Goal: Find

$$\begin{aligned} \max_{n_1,\dots,n_L} \quad & \sum_{k=1}^{K} \log M_k \\ \text{s.t.} \quad & \mathbb{E}\left[\tau\right] \leq N \\ & \mathbb{P}\left[(\hat{W}_1,\dots,\hat{W}_K) \neq (W_1,\dots,W_K)\right] \leq \epsilon \end{aligned}$$

Theorem 5 (Achievability)

There exists a VLSF code with L decoding times for the discrete memoryless MAC satisfying

$$\sum_{k=1}^{K} \log M_k = \frac{N\, I_K}{1-\epsilon} - \sqrt{N \log_{(L-1)}(N) \frac{V_K}{1-\epsilon}} + o(\sqrt{N})$$

$I_K = I(X_1,\dots,X_K;Y)$, $V_K = \mathrm{Var}\left[\imath(X_1,\dots,X_K;Y)\right]$

- Proof: We employ a single information density threshold decoder $\imath(X_{[K]}^{n_\ell}(m_{[K]}); Y^{n_\ell}) \geq \gamma$.

Summary of Part 2

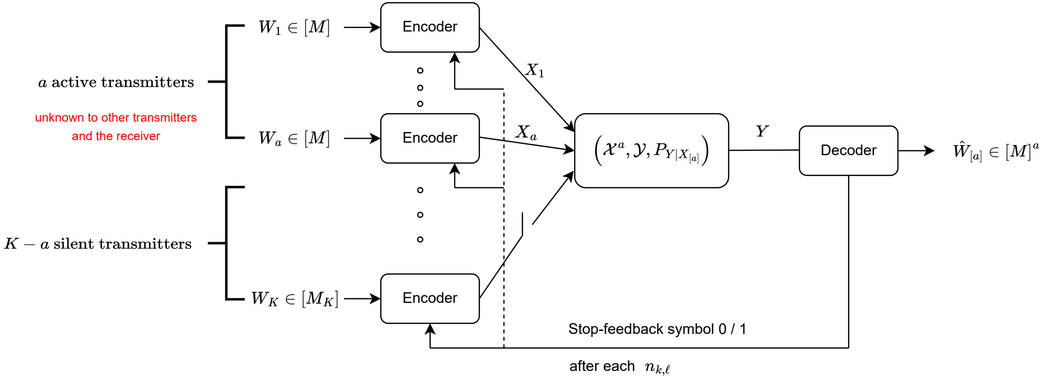
- Second-order achievability bounds for sparse VLSF codes over point-to-point and multiple access channels

- Optimizing the placements of L decoding times is critical to achieve high rates

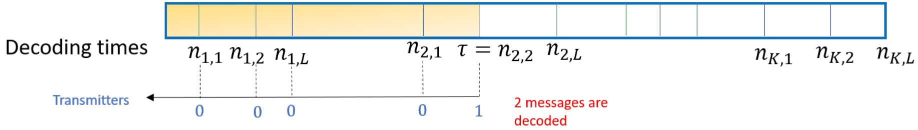
- A handful of decoding times is almost as good as decoding after every symbol
  For the BSC(0.11) at $N = 1000$ and $\epsilon = 10^{-3}$:
  - $L = 4$ achieves 97.0% of the rate achieved by $L = \infty$
  - $L = 1$ achieves 84.2% of the rate achieved by $L = \infty$

Outline

Rateless RAC (Random Access Channel) Communication that Uses Stop-Feedback



- **RAC**: A family of MACs up to K transmitters, any $a \leq K$ transmitters can be active
- **Compound channel model:** No probability of being active is assigned to transmitters
- **Agnostic channel model:** Nobody knows who are active
- We extend VLSF codes to the RAC
- Available decoding times $= n_{a,1}, \ldots, n_{a,L}$ for decoding $a \in [K]$ messages



Goal: Find

$$\max_{n_{a,\ell}\,:\, a\in[K],\ell\in[L]} M$$
$$\text{s.t.} \quad \mathbb{E}\left[\tau_a\right] \leq N_a \qquad a \in [K]$$
$$\mathbb{P}\left[(\hat{W}_1, \ldots, \hat{W}_a) \neq (W_1, \ldots, W_a)\right] \leq \epsilon_a \qquad a \in [K]$$

- Presented in the candidacy talk.

Theorem 6 (Achievability)

Fix an input distribution P_X. For a RAC satisfying some mild symmetry conditions, there exists a RAC code with K transmitters provided that

$$a \log M \leq N_a I_a - \sqrt{N_a V_a} Q^{-1}(\epsilon_a) - \frac{1}{2} \log N_a + O(1) \qquad \forall\, a \in [K]$$

We achieve the same first- and second-order terms as the best-known codes (e.g., [Tan-Kosut 14'], [Scarlett et al. 15]) for the MAC in operation.

- Proof: At times $n_{1,1}, \ldots, n_{K,1}$, we use information density threshold rule

- Presented in the candidacy talk.

Theorem 7

There exists a RAC code for the Gaussian RAC with K transmitters provided that

$$a \log M \leq N_a C(aP) - \sqrt{N_a V_a(P)} Q^{-1}(\epsilon_a) + \frac{1}{2} \log N_a + O(1) \qquad \forall\ a \in [K]$$

where

$$C(P) = \frac{1}{2} \log(1+P), \quad V_a(P) = \frac{(2a^2 - a)P^2 + 2aP}{2(aP+1)^2}$$

- Proof analyzes the error probability of the random code where codewords are generated uniformly on the restricted power sphere

Theorem 8 (Achievability)

Fix K, $L \geq 2$, $\epsilon \in (0,1)$, and a distribution P_X. Under some mild symmetry conditions, there exists a VLSF code for the discrete memoryless RAC with L decoding times for each $a \in [K]$ provided that

$$a \log M \leq \frac{N_a I_a}{1-\epsilon_a} - \sqrt{N_a \log_{(L-1)}(N_a) \frac{V_a}{1-\epsilon_a}} + o(\sqrt{N_a}) \qquad \forall\, a \in [K]$$

We achieve the same first- and second-order terms as the MAC in operation

- **Decoder:** At time n_0, the decoder applies a multiple hypothesis test to estimate # of active transmitters a

  If $\hat{a}$ is decoder's estimate, decoder uses a VLSF code for the $\hat{a}$-MAC with decoding times $\{n_{\hat{a},1}, \ldots, n_{\hat{a},L}\}$

Outline

Take-home message: Part 1

- In the low-latency, high-reliability regime, our MD approximation is more accurate than several state-of-the-art approximations
- In this regime, channel skewness cannot be neglected to get tight approximations

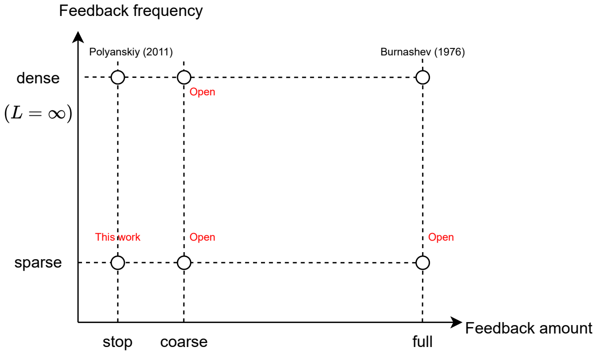
Take-home message: Part 2

- A handful of decoding times in VLSF codes achieve rates close to those achieved by $L = \infty$

Take-home message: Part 3

- For RACs, we achieve the same first- and second-order rates as if the active transmitters are known a priori
- Rateless coding with stop-feedback enables this result

- **Different ways of limiting the feedback in VLSF codes:**



  - Coarse feedback: Send R_f bits of feedback at each time
  - In many applications, we can send bursty feedback

- **A tight converse for VLSF codes with L decoding times** is missing

- Towards this goal, we have a non-asymptotic converse bound based on fundamental limits of SHTs
  This bound seems to be very challenging to analyze

- **Second-order converse for the RAC result:**
  - A second-order converse result for the K-MAC is also a converse result for the RAC
  - However, it is a long-standing open problem
  - [Kosut 22'] proves that the second-order term scales as $-O(\sqrt{N})$

$$K \log M^* = n_K I_K - O(\sqrt{n_K})$$

Thank you!

1. R. C. Yavas, V. Kostina, and M. Effros, “Third-order analysis of channel coding in the moderate deviations regime,” To be submitted to IT Transactions, Aug. 2022.
2. ——, “Third-order analysis of channel coding in the moderate deviations regime,” in 2022 IEEE International Symposium on Information Theory (ISIT), June 2022.
3. ——, “Variable-length sparse feedback codes for point-to-point, multiple access and random access channels,” To be submitted to IT Transactions, Aug. 2022.
4. ——, “Gaussian multiple and random access channels: Finite-blocklength analysis,” IEEE Transactions on Information Theory, vol. 67, no. 11, pp. 6983–7009, Nov. 2021
5. ——, “Nested sparse feedback codes for point-to-point, multiple access, and random access channels,” in 2021 IEEE Information Theory Workshop (ITW), Oct. 2021, pp. 1–6.
6. ——, “Random access channel coding in the finite blocklength regime,” IEEE Transactions on Information Theory, vol. 67, no. 4, pp. 2115– 2140, Apr. 2021.
7. ——, “Variable-length feedback codes with several decoding times for the gaussian channel,” in 2021 IEEE International Symposium on Information Theory (ISIT), Jun. 2021, pp. 1883–1888.
8. ——, “Gaussian multiple and random access in the finite blocklength regime,” in 2020 IEEE International Symposium on Information Theory (ISIT), June 2020, pp. 3013–3018.
9. M. Effros, V. Kostina, and R. C. Yavas, “Random access channel coding in the finite blocklength regime,” in 2018 IEEE International Symposium on Information Theory (ISIT), June 2018, pp. 1261–1265.

References

1. Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2307–2359, May 2010.
2. V. Strassen, "Asymptotische abschätzugen in Shannon's informationstheorie," in Trans. Third Prague Conf. Inf. Theory, Prague, 1962, pp. 689–723.
3. Y. Polyanskiy, H. V. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," IEEE Trans. Inf. Theory, vol. 57, no. 8, pp. 4903–4925, Aug. 2011.
4. O. Kosut, "A second-order converse bound for the multiple-access channel via wringing dependence," IEEE Tran, Inf. Theory, vol. 68, no. 6, pp. 3552–3584, Jun. 2022.
5. V. V. Petrov, Sums of independent random variables. New York, USA: Springer, Berlin, Heidelberg, 1975.
6. P. Moulin, "The log-volume of optimal codes for memoryless channels, asymptotically within a few nats," IEEE Trans. Inf. Theory, vol. 63, no. 4, pp. 2278–2313, Apr. 2017.
7. Y. Altuğ and A. B. Wagner, "Refinement of the random coding bound," IEEE Trans. Inf. Theory, vol. 60, no. 10, pp. 6005–6023, Oct. 2014.
8. Y. Altuğ and A. B. Wagner, "Moderate deviations in channel coding," IEEE Trans. Inf. Theory, vol. 60, no. 8, pp. 4417–4426, Aug. 2014.
9. M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," Problems of Information Transmission, vol. 12, no. 4, pp. 10–30, 1976.
10. K. Vakilinia, S. V. S. Ranganathan, D. Divsalar, and R. D. Wesel, "Optimizing transmission lengths for limited feedback with nonbinary ldpc examples," IEEE Transactions on Communications, vol. 64, no. 6, pp. 2245–2257, 2016.