Variable-Length Stop-Feedback Codes

Guodong SUN

Inria Centre d'Université Côte d'Azur

Advisors: Samir M. Perlaza, Philippe Mary, and Jean-Marie Gorce

May 21, 2026

Outline

Variable-Length Stop-Feedback (VLSF) Codes

Optimization of Sparse VLSF Codes via Saddlepoint Approximation

VLSF Codes over Correlated Noncoherent Fading Channels

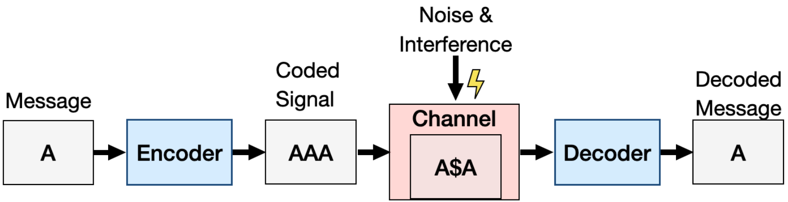
Outline

Variable-Length Stop-Feedback (VLSF) Codes

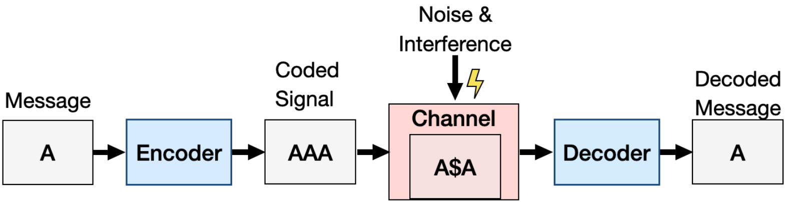
Optimization of Sparse VLSF Codes via Saddlepoint Approximation

VLSF Codes over Correlated Noncoherent Fading Channels

Fixed-length Channel Coding



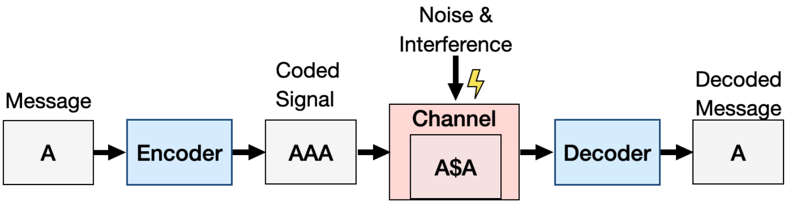
Fixed-length Channel Coding



Examples

- LDPC codes: Wi-Fi, Cellular Networks (4G, 5G).

Fixed-length Channel Coding



Examples

- LDPC codes: Wi-Fi, Cellular Networks (4G, 5G).
- Turbo codes: used in 3G and 4G systems.

A Tale of Two Conversations

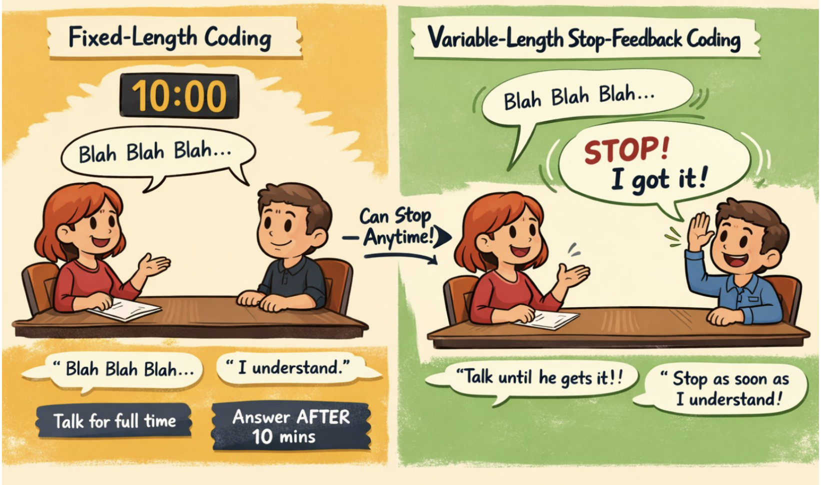


Image generated by AI from my story.

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$
  - Message m is mapped to $f(m) = x_m^\infty$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$
  - Message m is mapped to $f(m) = x_m^\infty$
- Over a noisy channel, the output is random.
  - Output $Y \sim P_{Y^\infty | X^\infty = x_m^\infty}$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$
  - Message m is mapped to $f(m) = x_m^\infty$
- Over a noisy channel, the output is random.
  - Output $Y \sim P_{Y^\infty|X^\infty=x_m^\infty}$
- Decoder returns $\hat{W} = g(Y^\tau)$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$
  - Message m is mapped to $f(m) = x_m^\infty$
- Over a noisy channel, the output is random.
  - Output $Y \sim P_{Y^\infty | X^\infty = x_m^\infty}$
- Decoder returns $\hat{W} = g(Y^\tau)$
  - Decoding stops at $\tau \triangleq \inf\{n : \mathbb{P}(\hat{W} \neq m | Y^\tau) \leq \epsilon\}$

System Model

$$W \xrightarrow{f} X^\infty \xrightarrow{P_{Y|X}} Y^\infty \xrightarrow{g} \hat{W}$$

- Message $W = m \in \{1, 2, \cdots M\}$
- Codebook $\mathcal{C} \triangleq \{x_1^\infty, x_2^\infty, \cdots x_M^\infty\}$
  - Random coding: Each codeword $x^\infty \overset{\text{i.i.d.}}{\sim} P_X^\infty$
  - Message m is mapped to $f(m) = x_m^\infty$
- Over a noisy channel, the output is random.
  - Output $Y \sim P_{Y^\infty|X^\infty = x_m^\infty}$
- Decoder returns $\hat{W} = g(Y^\tau)$
  - Decoding stops at $\tau \triangleq \inf\{n : \mathbb{P}(\hat{W} \neq m|Y^\tau) \leq \epsilon\}$

Definition: VLSF codes

A (M, ϵ) VSLF code achieves a message size M with average error probability $\mathbb{P}(\hat{W} \neq W) \leq \epsilon$.

Hypothesis Testing

Binary Hypothesis Testing

$$(X^n, Y^n) \sim \begin{cases} \mathcal{H}_0 : P_{X^n} P_{Y^n|X^n} & (X^n \text{ was transmitted}) \\ \mathcal{H}_1 : P_{X^n} P_{Y^n} & (X^n \text{ was not transmitted}) \end{cases}$$

Hypothesis Testing

Binary Hypothesis Testing

$$(X^n, Y^n) \sim \begin{cases} \mathcal{H}_0 : P_{X^n} P_{Y^n|X^n} & (X^n \text{ was transmitted}) \\ \mathcal{H}_1 : P_{X^n} P_{Y^n} & (X^n \text{ was not transmitted}) \end{cases}$$

Information density: the log-likelihood ratio:

$$\imath(x^n; y^n) \triangleq \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$$

where P_{Y^n} is the marginal distribution induced by the input P_X^n.

Information Density Random Walks

Random coding + memoryless i.i.d. channels = random walks

$$\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$$

Information Density Random Walks

Random coding + memoryless i.i.d. channels = random walks

$$\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$$
$$= \sum_{i=1}^{n} \log \frac{\mathrm{d}P_{Y_i|X_i=x_i}}{\mathrm{d}P_{Y_i}}(y_i)$$

Information Density Random Walks

Random coding + memoryless i.i.d. channels = random walks

$$\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$$

$$= \sum_{i=1}^{n} \log \frac{\mathrm{d}P_{Y_i|X_i=x_i}}{\mathrm{d}P_{Y_i}}(y_i)$$

$$= \sum_{i=1}^{n} \imath(x_i; y_i)$$

Information Density Random Walks

Random coding + memoryless i.i.d. channels = random walks

$$\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$$
$$= \sum_{i=1}^{n} \log \frac{\mathrm{d}P_{Y_i|X_i=x_i}}{\mathrm{d}P_{Y_i}}(y_i)$$
$$= \sum_{i=1}^{n} \imath(x_i; y_i)$$

Characterizing VLSF codes via this **random-walk** structure.

Decoder

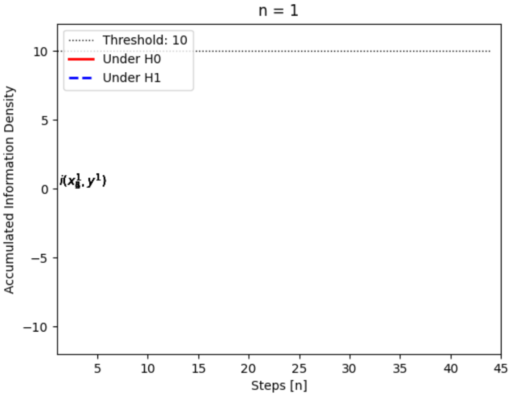
At each time n, the decoder tests M codewords by information densities:

$$\{\imath(x_1^n; y^n), \imath(x_2^n; y^n), \cdots, \imath(x_M^n; y^n)\}$$

Decoder

At each time n, the decoder tests M codewords by information densities:

$$\{\imath(x_1^n; y^n), \imath(x_2^n; y^n), \cdots, \imath(x_M^n; y^n)\}$$



Decoding Rule

The VLSF decoder stops at the first time τ where one of these densities exceeds a threshold γ:

$$\tau = \inf\{n : \max_m \imath(x_m^n; y^n) \geq \gamma\}$$

Decoding Rule

The VLSF decoder stops at the first time τ where one of these densities exceeds a threshold γ:

$$\tau = \inf\{n : \max_m \imath(x_m^n; y^n) \geq \gamma\}$$

Interpretation

Decoding stops when the target reliability is guaranteed.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ *and* $M \in \mathbb{N}_{>0}$. *For any threshold* $\gamma \geq \log \frac{M-1}{\epsilon}$, *the threshold-based VLSF decoding rule satisfies*

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

---

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ *and* $M \in \mathbb{N}_{>0}$*. For any threshold* $\gamma \geq \log \frac{M-1}{\epsilon}$*, the threshold-based VLSF decoding rule satisfies*

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$

---

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
   - Non-transmitted codeword $(x^n; y^n) \sim P_{X^n} P_{Y^n}$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
   - Non-transmitted codeword $(x^n; y^n) \sim P_{X^n} P_{Y^n}$
   - Nonnegative martingale with $\mathbb{E}[\mathsf{L}_n] = 1$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
   - Non-transmitted codeword $(x^n; y^n) \sim P_{X^n} P_{Y^n}$
   - Nonnegative martingale with $\mathbb{E}[\mathsf{L}_n] = 1$
3. Ville's inequality[1]: $\mathbb{P}\left(\sup_{n \geq 1} \mathsf{L}_n \geq e^{\gamma}\right) \leq e^{-\gamma}$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
   - Non-transmitted codeword $(x^n; y^n) \sim P_{X^n}P_{Y^n}$
   - Nonnegative martingale with $\mathbb{E}[\mathsf{L}_n] = 1$
3. Ville's inequality[1]: $\mathbb{P}\left(\sup_{n\geq 1} \mathsf{L}_n \geq e^{\gamma}\right) \leq e^{-\gamma}$
4. By the union bound over the $M-1$ incorrect codewords, $\mathbb{P}[W \neq \hat{W}] \leq (M-1)e^{-\gamma}$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Reliability Guarantee

Theorem

Let $\epsilon > 0$ and $M \in \mathbb{N}_{>0}$. For any threshold $\gamma \geq \log \frac{M-1}{\epsilon}$, the threshold-based VLSF decoding rule satisfies

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon.$$

Proof.

1. Information density: $\imath(x^n; y^n) = \log \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
2. Likelihood ratio $\mathsf{L}_n \triangleq e^{\imath(x^n;y^n)} = \frac{\mathrm{d}P_{Y^n|X^n=x^n}}{\mathrm{d}P_{Y^n}}(y^n)$
   - Non-transmitted codeword $(x^n; y^n) \sim P_{X^n}P_{Y^n}$
   - Nonnegative martingale with $\mathbb{E}[\mathsf{L}_n] = 1$
3. Ville's inequality[1]: $\mathbb{P}\left(\sup_{n\geq 1} \mathsf{L}_n \geq e^{\gamma}\right) \leq e^{-\gamma}$
4. By the union bound over the $M-1$ incorrect codewords, $\mathbb{P}[W \neq \hat{W}] \leq (M-1)e^{-\gamma}$
5. Choosing $\gamma \geq \log \frac{M-1}{\epsilon}$ yields $\mathbb{P}[W \neq \hat{W}] \leq \epsilon$

$\square$

[1] J. Ville, *Etude critique de la notion de collectif*. Gauthier-Villars, first ed., 1939.

Latency:

Stopping time τ:

- τ is upper-bounded by the true message (indexed by 1).

[2] W. Feller, *An introduction to probability theory and its applications, Volume 2*.John Wiley & Sons, 1991.

Latency:

Stopping time τ:

- τ is upper-bounded by the true message (indexed by 1).
- **Average latency:** $\mathbb{E}[\tau] \leq \frac{\gamma}{\mathbb{E}_{\mathcal{H}_0}[\imath(X_1;Y)]}$

[2] W. Feller, *An introduction to probability theory and its applications, Volume 2*.John Wiley & Sons, 1991.

Latency:

Stopping time τ:

- τ is upper-bounded by the true message (indexed by 1).
- **Average latency:** $\mathbb{E}[\tau] \leq \frac{\gamma}{\mathbb{E}_{\mathcal{H}_0}[\imath(X_1;Y)]}$
  - Wald's identity[2]

[2] W. Feller, *An introduction to probability theory and its applications, Volume 2*.John Wiley & Sons, 1991.

Latency:

Stopping time τ:

- τ is upper-bounded by the true message (indexed by 1).
- **Average latency:** $\mathbb{E}[\tau] \leq \frac{\gamma}{\mathbb{E}_{\mathcal{H}_0}[\imath(X_1;Y)]}$
  - Wald's identity[2]
    1. Suppose $Z_1, Z_2, \cdots$ are i.i.d. random variables, $Z_i \sim P_Z$
    2. Let T be a stopping time for $(Z_i)_{i\geq 1}$.
    3. Suppose that $\mathbb{E}[T], \mathbb{E}[|Z|] < \infty$.
    4. Then:

$$\mathbb{E}\left[\sum_{i=1}^{T} Z_i\right] = \mathbb{E}[T]\mathbb{E}[Z]$$

[2] W. Feller, *An introduction to probability theory and its applications, Volume 2*.John Wiley & Sons, 1991.

Latency:

Stopping time τ:

- τ is upper-bounded by the true message (indexed by 1).
- **Average latency:** $\mathbb{E}[\tau] \leq \frac{\gamma}{\mathbb{E}_{\mathcal{H}_0}[\imath(X_1;Y)]}$
  - Wald's identity[2]
    1. Suppose $Z_1, Z_2, \cdots$ are i.i.d. random variables, $Z_i \sim P_Z$
    2. Let T be a stopping time for $(Z_i)_{i\geq 1}$.
    3. Suppose that $\mathbb{E}[T], \mathbb{E}[|Z|] < \infty$.
    4. Then:

$$\mathbb{E}\left[\sum_{i=1}^{T} Z_i\right] = \mathbb{E}[T]\mathbb{E}[Z]$$

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i=1}^{T} Z_i\right] \Big/ \mathbb{E}[Z]$$

[2]W. Feller, *An introduction to probability theory and its applications, Volume 2*.John Wiley & Sons, 1991.

Approximation of stopping time distribution[3]

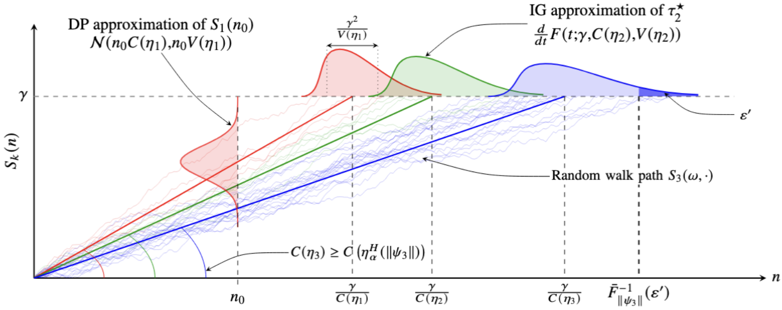


Figure: Diffusion processes (DP) approximation and inverse Gaussian (IG) distributed stopping time under different SNRs

[3] Joint work with Dr. Kevin Zagalo

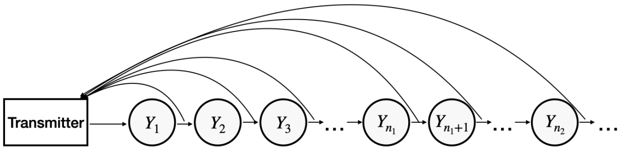
Outline

Variable-Length Stop-Feedback (VLSF) Codes

Optimization of Sparse VLSF Codes via Saddlepoint Approximation

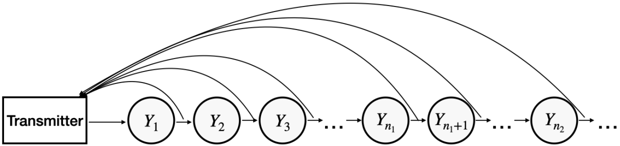
VLSF Codes over Correlated Noncoherent Fading Channels

Observation

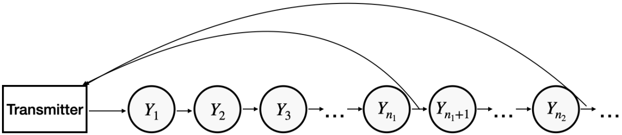


Dense feedback: Feedback after every symbol: n symbols $=$ n feedback

Observation



Dense feedback: Feedback after every symbol: n symbols $=$ n feedback



Sparse feedback: Feedback occurs only at t **optimized** time instants: n symbols $=$ t feedback

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j) \mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]: Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants: $\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols
  - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants: $\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols
  - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols
  - With $\mathbb{P}[\imath(X_1^{n_2}, Y^{n_2}) < \gamma]$, transmit $(n_3 - n_2)$ symbols
  - ...

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]: Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols
  - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols
  - With $\mathbb{P}[\imath(X_1^{n_2}, Y^{n_2}) < \gamma]$, transmit $(n_3 - n_2)$ symbols
  - ...
  - Continue until n_t

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols
  - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols
  - With $\mathbb{P}[\imath(X_1^{n_2}, Y^{n_2}) < \gamma]$, transmit $(n_3 - n_2)$ symbols
  - ...
  - Continue until n_t
- Constraints:

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
  - First transmit n_1 symbols
  - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols
  - With $\mathbb{P}[\imath(X_1^{n_2}, Y^{n_2}) < \gamma]$, transmit $(n_3 - n_2)$ symbols
  - ...
  - Continue until n_t
- Constraints:
  - Before n_t: Threshold-rule error guarantee $(M-1)e^{-\gamma}$

[4] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem:

Literature[4]**:** Define t decoding time instants:
$\mathcal{T} \triangleq \{n_1, n_2, \ldots, n_t\}$

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{1a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{1b}$$

- Objective: Average decoding time $\mathbb{E}[\tau]$:
    - First transmit n_1 symbols
    - With $\mathbb{P}[\imath(X_1^{n_1}, Y^{n_1}) < \gamma]$, transmit $(n_2 - n_1)$ symbols
    - With $\mathbb{P}[\imath(X_1^{n_2}, Y^{n_2}) < \gamma]$, transmit $(n_3 - n_2)$ symbols
    - ...
    - Continue until n_t
- Constraints:
    - Before n_t: Threshold-rule error guarantee $(M-1)e^{-\gamma}$
    - At n_t: No crossing: $\mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma]$

[4] 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," in *2022 ISIT*, pp. 1527–1532

Proposed rule

If no threshold-crossing occurs at n_t, the decoder selects the codeword with the highest information density:

- Error probability characterized by fixed blocklength analysis[5]

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{2a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \epsilon_{fb}^*(M, n_t) \leq \epsilon \tag{2b}$$

[5] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

Proposed rule

If no threshold-crossing occurs at n_t, the decoder selects the codeword with the highest information density:

- Error probability characterized by fixed blocklength analysis[5]

$$\min_{\gamma, \mathcal{T}} \quad n_1 + \sum_{j=1}^{|\mathcal{T}|-1} (n_{j+1} - n_j)\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma], \tag{2a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \epsilon^*_{fb}(M, n_t) \leq \epsilon \tag{2b}$$

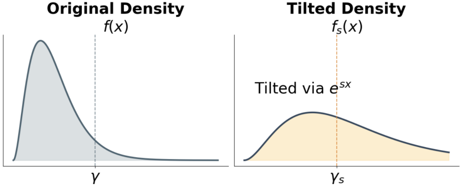
Unknown quantity

$$\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$$

[5] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

Saddlepoint Approximations

Step 1: $S_n = \sum_{i=1}^{n} \imath(X_{1,i}; Y_i)$ Exponential tilting

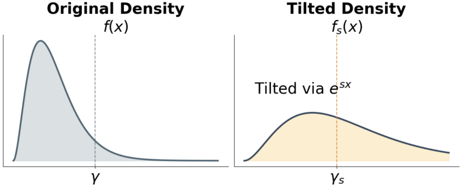


$$\boxed{f_s(x) = \frac{f(x)e^{sx}}{M(s)}}$$

$M(s) = \mathbb{E}[e^{s \cdot S_n}]$
Moment Generating Function

Saddlepoint Approximations

Step 1: $S_n = \sum_{i=1}^{n} \imath(X_{1,i}; Y_i)$ Exponential tilting



$$\boxed{f_s(x) = \frac{f(x)e^{sx}}{M(s)}}$$

$M(s) = \mathbb{E}[e^{s \cdot S_n}]$
Moment Generating Function

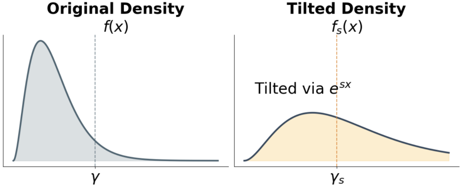
Step 2: Find **Saddlepoint** $\hat{s}$ by solving

$$\boxed{K'(s) = \frac{\gamma}{n}}$$

$K(s) = \log M(s)$
Cumulant Generating Function

Saddlepoint Approximations

Step 1: $S_n = \sum_{i=1}^{n} \imath(X_{1,i}; Y_i)$ Exponential tilting



$$f_s(x) = \frac{f(x)e^{sx}}{M(s)}$$

$M(s) = \mathbb{E}[e^{s \cdot S_n}]$
Moment Generating Function

Step 2: Find **Saddlepoint** $\hat{s}$ by solving

$$K'(s) = \frac{\gamma}{n}$$

$K(s) = \log M(s)$
Cumulant Generating Function

Step 3: The Lugannani-Rice Approximation

$$\mathbb{P}[S_n \leq \gamma] \approx \Phi(\hat{w}) + \phi(\hat{w})\left(\frac{1}{\hat{w}} - \frac{1}{\hat{u}}\right)$$

$\Phi(\,\cdot\,)$: Standard Normal CDF
$\phi(\,\cdot\,)$: Standard Normal PDF

Definition of terms:

$\hat{w} = \mathrm{sgn}(\hat{s})\sqrt{2(\gamma\hat{s} - nK(\hat{s}))}$:
Signed root of log-likelihood ratio deviation

$\hat{u} = \hat{s}\sqrt{nK''(\hat{s})}$
Related to curvature of tilted distribution

Saddlepoint Approximations

1. $\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$ admits a closed-form expression in terms of the saddlepoint $\hat{s}$

Saddlepoint Approximations

1. $\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$ admits a closed-form expression in terms of the saddlepoint $\hat{s}$
2. Saddlepoint $\hat{s}$ is in closed form with respect to γ and n for common channels

Saddlepoint Approximations

1. $\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$ admits a closed-form expression in terms of the saddlepoint $\hat{s}$
2. Saddlepoint $\hat{s}$ is in closed form with respect to γ and n for common channels

Results

Optimization problem can be solved analytically.

Accuracy of Saddlepoint Approximations

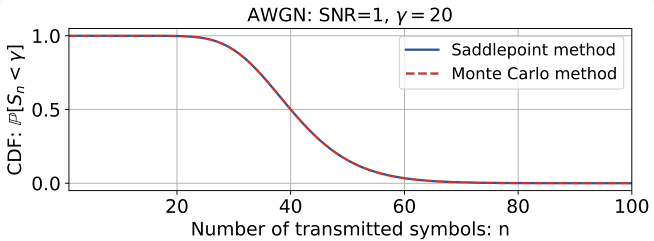


Figure: Saddlepoint approximations for a continuous channel

Accuracy of Saddlepoint Approximations

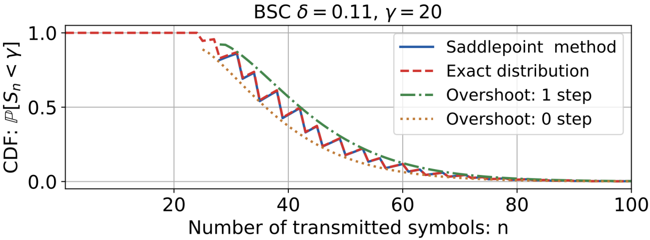
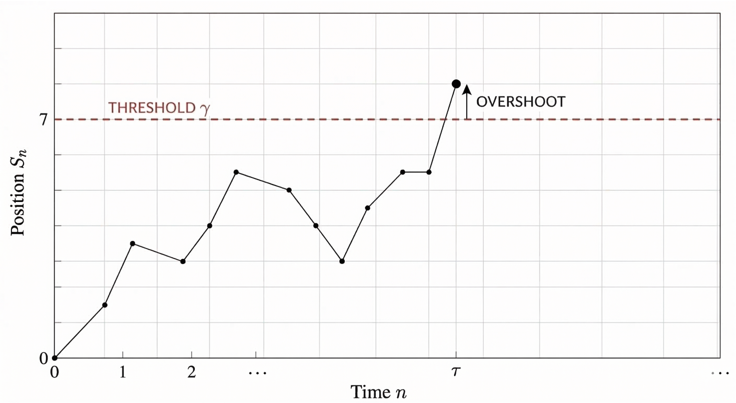
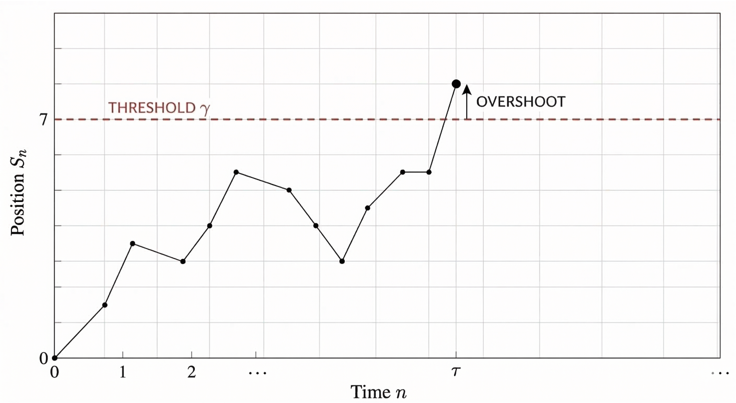


Figure: Saddlepoint approximations for a discrete channel

Random Walk Overshoot



Random Walk Overshoot



Assuming an overshoot of 0 or 1 step yields continuous correction bounds that can be used for gradient-based optimization.

Accuracy of Saddlepoint Approximations

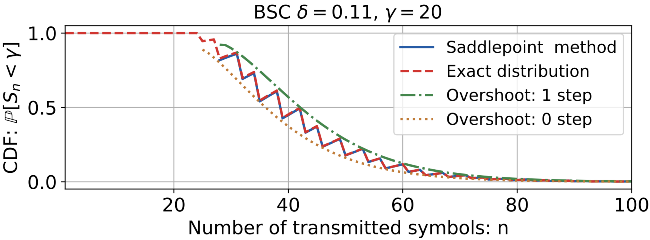


Figure: Saddlepoint approximations for discrete channel.

Optimization problem

Literature[6]: No crossing at final decoding: = error

$$\min_{\gamma, \mathcal{T}} \quad \mathbb{E}[\tau] \tag{3a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{3b}$$

[6] 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," in *2022 ISIT*, pp. 1527–1532

Optimization problem

Literature[6]**:** No crossing at final decoding: $=$ error

$$\min_{\gamma, \mathcal{T}} \quad \mathbb{E}[\tau] \tag{3a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \mathbb{P}[\imath(X_1^{n_t}, Y^{n_t}) < \gamma] \leq \epsilon. \tag{3b}$$

Proposed rule: No crossing at final decoding $=$ select the highest information density

$$\min_{\gamma, \mathcal{T}} \quad \mathbb{E}[\tau] \tag{4a}$$

$$\text{s.t.} \quad (M-1)e^{-\gamma} + \epsilon^*_{fb}(M, n_t) \leq \epsilon \tag{4b}$$

[6] 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," in *2022 ISIT*, pp. 1527–1532

Results: Number of Feedback and Achievable Rates

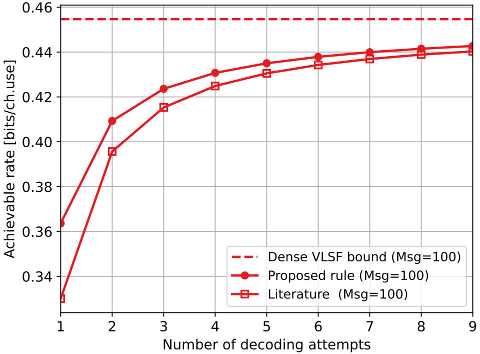


Figure: Achievable rate of sparse VLSF codes versus number of decoding attempts.

Results: Number of Feedback and Achievable Rates

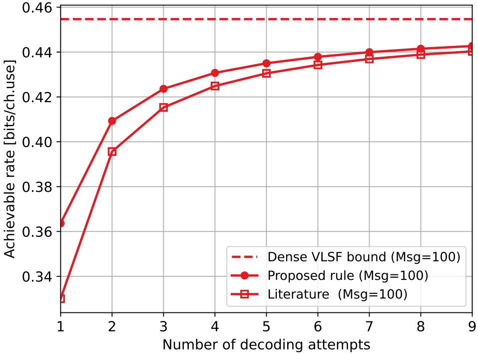


Figure: Achievable rate of sparse VLSF codes versus number of decoding attempts.

- Near-optimal with a small number of decoding attempts.

Results: Number of Feedback and Achievable Rates

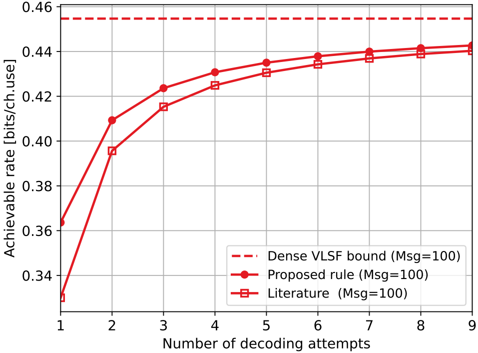


Figure: Achievable rate of sparse VLSF codes versus number of decoding attempts.

- Near-optimal with a small number of decoding attempts.
- Refined decoding rule yields large gains for one or two attempts.

Results: Binary Symmetric Channel

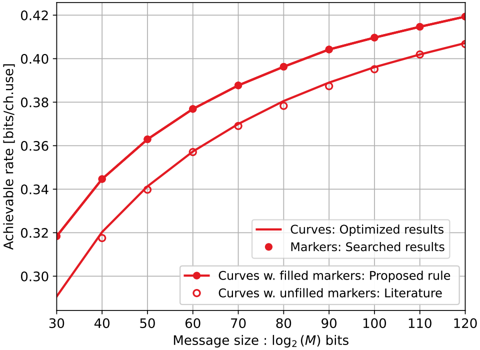


Figure: Achievable rate versus message size (in bits)

- BSC ($p = 0.11$, $t = 3$ attempts)

Results: Binary Symmetric Channel

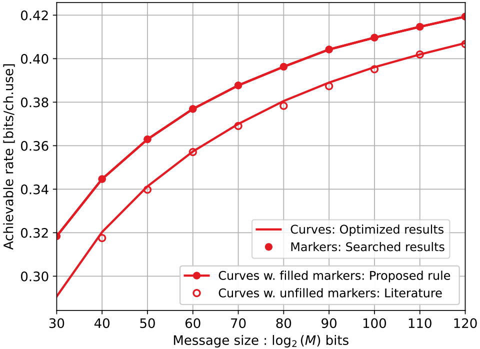


Figure: Achievable rate versus message size (in bits)

- BSC ($p = 0.11$, $t = 3$ attempts)
- Proposed rule achieves higher rates

Results: Binary Symmetric Channel

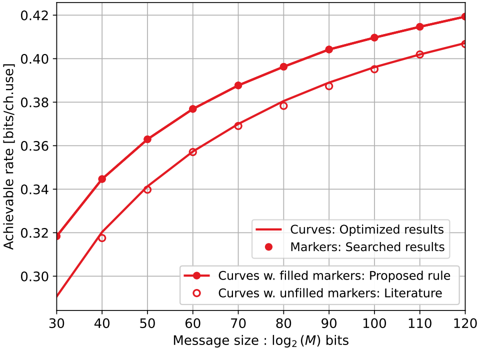


Figure: Achievable rate versus message size (in bits)

- BSC ($p = 0.11$, $t = 3$ attempts)
- Proposed rule achieves higher rates
- Near optimal rate

Results: BSC

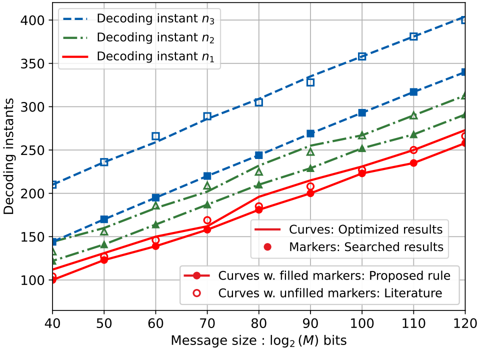


Figure: Optimal decoding instants versus message size (in bits)

- BSC ($p = 0.11$, $t = 3$ attempts)

Results: BSC

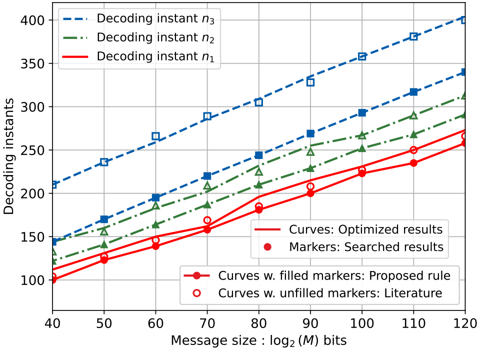


Figure: Optimal decoding instants versus message size (in bits)

- BSC ($p = 0.11$, $t = 3$ attempts)
- Optimized decoding instants occur earlier at later stages ($n2$, $n3$)

Outline

Variable-Length Stop-Feedback (VLSF) Codes

Optimization of Sparse VLSF Codes via Saddlepoint Approximation

VLSF Codes over Correlated Noncoherent Fading Channels

Correlated Noncoherent Fading Channels

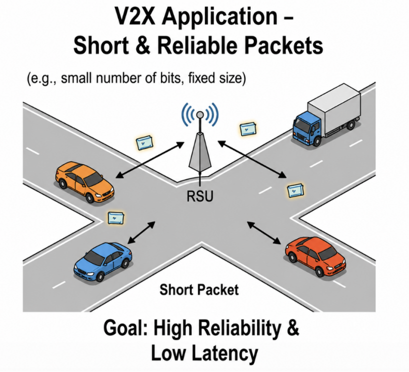


Figure: Vehicular Networks; RSU: Roadside Units

Correlated Noncoherent Fading Channels

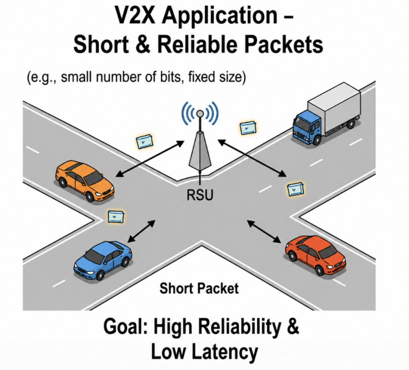


Figure: Vehicular Networks; RSU: Roadside Units

Model Dilemma:

- Fast fading: dynamic channel over codeword duration
- Pilots: channel outdated quickly
- Correlated Noncoherent Fading Channels

Correlated Noncoherent Fading Channels

"Middle Ground" regime

Channels experience symbol-wise correlation, but not block-fading.
Fading channel with memory[7]

[7] Lapidoth, Amos, and Stefan M. Moser. "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels." IEEE Transactions on Information Theory 49.10 (2003): 2426-2467.

Correlated Noncoherent Fading Channels

"Middle Ground" regime

Channels experience symbol-wise correlation, but not block-fading.
Fading channel with memory[7]

Example: First-order Gauss-Markov process with correlation ρ

$$H_k = \rho H_{k-1} + \sqrt{1-\rho^2} W_k,$$

[7] Lapidoth, Amos, and Stefan M. Moser. "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels." IEEE Transactions on Information Theory 49.10 (2003): 2426-2467.

Martingale Structure

Analysis based on martingale for any nontransmitted codeword is still valid!

Martingale Structure

Analysis based on martingale for any nontransmitted codeword is still valid!

Conditional expectation with respect to $\mathcal{F}_{n-1}$ under $\mathcal{H}_1$:

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{H}_1}[\mathsf{L}_n(X^n; Y^n)|\mathcal{F}_{n-1}] \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})\mathbb{E}_{\mathcal{H}_1}\left[\frac{\mathrm{d}P_{Y_n|X_n,Y^{n-1}}}{\mathrm{d}P_{Y_n|Y^{n-1}}}(Y_n)|\mathcal{F}_{n-1}\right] \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})\int \frac{\mathrm{d}P_{Y_n|X_n,Y^{n-1}}}{\mathrm{d}P_{Y_n|Y^{n-1}}}(Y_n)\mathrm{d}P_{Y_n|Y^{n-1}}(Y_n) \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})
\end{aligned}
$$

Martingale Structure

Analysis based on martingale for any nontransmitted codeword is still valid!

Conditional expectation with respect to $\mathcal{F}_{n-1}$ under $\mathcal{H}_1$:

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{H}_1}[\mathsf{L}_n(X^n; Y^n)|\mathcal{F}_{n-1}] \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})\mathbb{E}_{\mathcal{H}_1}\left[\frac{\mathrm{d}P_{Y_n|X_n,Y^{n-1}}}{\mathrm{d}P_{Y_n|Y^{n-1}}}(Y_n)|\mathcal{F}_{n-1}\right] \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})\int \frac{\mathrm{d}P_{Y_n|X_n,Y^{n-1}}}{\mathrm{d}P_{Y_n|Y^{n-1}}}(Y_n)\mathrm{d}P_{Y_n|Y^{n-1}}(Y_n) \\
=&\mathsf{L}_{n-1}(X^{n-1}; Y^{n-1})
\end{aligned}
$$

Takeaway

VLSF Decoding rule still holds:

- Decoder declares the message for which the information density crosses the threshold.
- Reliability Guarantee still holds.

Challenges

The actual information density is intractable for **correlated noncoherent** channel:

Challenges

The actual information density is intractable for **correlated noncoherent** channel:

For any measurable set $A \in \mathcal{B}(\mathcal{Y}^n)$,

$$\begin{aligned} P_{Y^n}(A) &= \int\int P_{Y^n|X^n=x^n,H^n=h^n}(A)\mathrm{d}P_{X^n}(x^n)\mathrm{d}P_{H^n}(h^n) \\ &= \int P_{Y^n|H^n=h^n}(A)\mathrm{d}P_{H^n}(h^n). \end{aligned}$$

Challenges

The actual information density is intractable for **correlated noncoherent** channel:

For any measurable set $A \in \mathcal{B}(\mathcal{Y}^n)$,

$$P_{Y^n}(A) = \int \int P_{Y^n|X^n=x^n, H^n=h^n}(A) \mathrm{d}P_{X^n}(x^n) \mathrm{d}P_{H^n}(h^n)$$
$$= \int P_{Y^n|H^n=h^n}(A) \mathrm{d}P_{H^n}(h^n).$$

Solutions

Change-of-measure + Variation approach

$$L(h^n) \triangleq \frac{\mathrm{d}P_{H^n}}{\mathrm{d}Q_{H^n}}(h^n).$$

where $\forall Q_{H^n} \in \mathcal{Q}$, and $\mathcal{Q}$ the set of reference probability measures (**computable**) such that $P_{H^n} \ll Q_{H^n}$.

Information density lower bound

Theorem (Information density lower bound)

Let $Q_{H^n} \in \mathcal{Q}$ be a reference measure. Fix $r > 1$, and let $s \triangleq \frac{r}{r-1}$ denote its Hölder conjugate. Then, given a pair $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$,

$$\begin{aligned} \imath(x^n; y^n) \geq \log f_{Y^n|X^n}(y^n|x^n) - \frac{r-1}{r} D_r(P_{H^n} \| Q_{H^n}) \\ - \frac{1}{s} \log \mathbb{E}_{Q_{H^n}} \left[f_{Y^n|H^n}(y^n|H^n)^s \right], \end{aligned} \tag{5}$$

where $D_r(P_{H^n} \| Q_{H^n})$ is the Rényi divergence of order r.

Information density lower bound

Theorem (Information density lower bound)

Let $Q_{H^n} \in \mathcal{Q}$ be a reference measure. Fix $r > 1$, and let $s \triangleq \frac{r}{r-1}$ denote its Hölder conjugate. Then, given a pair $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$,

$$\begin{aligned}\imath(x^n; y^n) \geq \log f_{Y^n|X^n}(y^n|x^n) - \frac{r-1}{r} D_r(P_{H^n} \| Q_{H^n}) \\ - \frac{1}{s} \log \mathbb{E}_{Q_{H^n}} \left[f_{Y^n|H^n}(y^n|H^n)^s \right],\end{aligned} \tag{5}$$

where $D_r(P_{H^n} \| Q_{H^n})$ is the Rényi divergence of order r.

Operational meaning

Whenever the lower bound crosses the threshold, the actual information density must cross, too. The VLSF decoding rule holds with reliability guarantee.

Upper bound

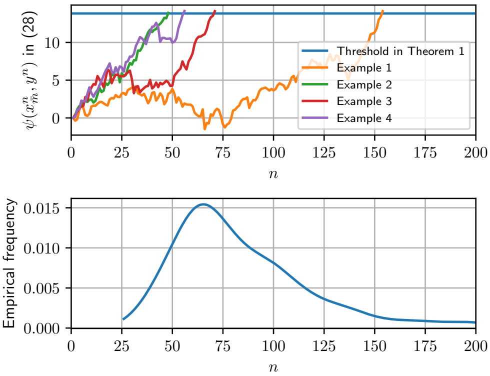
Theorem (Information density upper bound)

Let $Q_{H^n} \in \mathcal{Q}$*, and the likelihood ratio* $L(h^n)$*. Then, for every realization* (x^n, y^n)*,*

$$\imath(x^n; y^n) \leq \log f_{Y^n|X^n}(y^n|x^n) + D(P_{H^n} \| Q_{H^n}) - \mathbb{E}_{Q_{H^n}}\left[\log f_{Y^n|H^n}(y^n|H^n)\right], \tag{6}$$

where $D(P_{H^n} \| Q_{H^n})$ *denotes the Kullback-Leibler (KL) divergence.*

Performance characterization



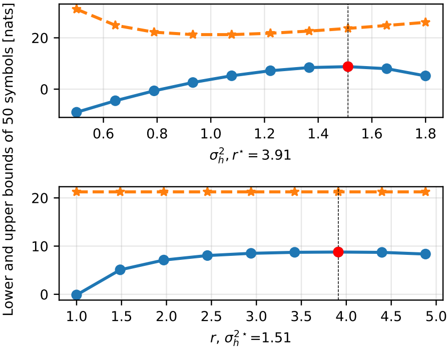
8

Specialization

- Q_{H^n} i.i.d. Gaussian auxiliary distribution
- P_X Gaussian signaling

[8] Chosen for mathematical tractability rather than strict optimality

Upper- and lower bound



- Non-negligible gap.

Conclusion

Conclusion

- We develop a gradient-based framework to compute achievability bounds for sparse VLSF codes, leveraging the analyticity of saddlepoint approximations.

Conclusion

- We develop a gradient-based framework to compute achievability bounds for sparse VLSF codes, leveraging the analyticity of saddlepoint approximations.
- The framework enables alternative decoding rules and reveals rate gains.

Conclusion

- We develop a gradient-based framework to compute achievability bounds for sparse VLSF codes, leveraging the analyticity of saddlepoint approximations.
- The framework enables alternative decoding rules and reveals rate gains.
- Furthermore, performance characterization of VLSF codes could be extended to correlated noncoherent fading channels.

References

R. W. Butler, *Saddlepoint approximations with applications*, vol. 22.
Cambridge University Press, 2007.

Backup Slides

Saddlepoint Approximations

- Let $Z_i = \imath(X_{1,i}; Y_i)$ be information density per channel use.
- Denote $K_Z(s)$ the cumulant generating function (CGF) of Z.
- The CGF of $S_n = \imath(X_1^n, Y^n) = \sum_{i=1}^n Z_i$ is $K(s) = nK_Z(s)$.
- If Z has a density, the Lugannani-Rice approximation [1] gives, for each $n \in \mathbb{N}_{>0}$ the CDF[9] of

$$\mathbb{P}[S_n < \gamma] \approx \begin{cases} \Phi(\hat{w}) + \phi(\hat{w})(1/\hat{w} - 1/\hat{u}) & \gamma \neq \mathbb{E}[S_n], \\ \frac{1}{2} + \frac{K'''(0)}{6\sqrt{2\pi}K''(0)^{3/2}} & \gamma = \mathbb{E}[S_n], \end{cases}$$

where $\hat{w} = \mathrm{sgn}(\hat{s})\sqrt{2(\hat{s}\gamma - K(\hat{s}))}$ and $\hat{u} = \hat{s}\sqrt{K''(\hat{s})}$.

Takeaway:

$\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$ admits a closed-form expression in terms of the saddlepoint $\hat{s}$.

[9] The functions ϕ and Φ denote the standard normal probability density and CDF, respectively, and sgn(.) is the sign function.

Saddlepoint Approximations

For discrete Z, let k denote the smallest attainable lattice point above γ. Then, $\mathbb{P}[S_n < \gamma] \approx$

$$\begin{cases} \Phi(\hat{w}) + \phi(\hat{w})(1/\hat{w} - 1/\bar{u}) & k \neq \mathbb{E}[S_n], \\ \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\left(\frac{K'''(0)}{6K''(0)^{3/2}} - \frac{1}{2\sqrt{K''(0)}}\right) & k = \mathbb{E}[S_n], \end{cases}$$

where $\hat{w}$ is defined as in the continuous case, and the curvature correction term is given by $\bar{u} = (1 - e^{-\hat{s}})\sqrt{K''(\hat{s})}$.

Takeaway:

$\mathbb{P}[\imath(X_1^{n_j}, Y^{n_j}) < \gamma]$ admits a closed-form expression in terms of the saddlepoint $\hat{s}$.

Common Memoryless i.i.d. Channels

Additive White Gaussian Noise (AWGN) Channel:
$Y = X + N$, where $X \sim \mathcal{N}(0, p_0)$ and $N \sim \mathcal{N}(0, 1)$

$$Z = \frac{1}{2}\log(1+p_0) + \frac{1}{2}\left(\frac{Y^2}{p_0+1} - (Y-X)^2\right).$$

Centering the random variable Z as $\tilde{Z} = Z - \frac{1}{2}\log(1+p_0)$ The threshold $\tilde{\gamma} = \gamma - \frac{n}{2}\log(1+p_0)$ The CGF of $\tilde{S}_n$ is given by

$$K_{\tilde{S}_n}(s) = -\frac{n}{2}\log\left(1 - \frac{p_0 s^2}{p_0+1}\right),$$

and the corresponding saddlepoint $\hat{s}$ is given by

$$\hat{s} = \frac{-n + \sqrt{n^2 + 4\tilde{\gamma}^2(p_0+1)/(p_0)}}{2\tilde{\gamma}}.$$

Takeaway:

Saddlepoint $\hat{s}$ is in closed form with respect to γ and n for common channels.

Common Memoryless i.i.d. Channels

Binary Symmetric Channel (BSC): with crossover probability $\delta \in (0, \frac{1}{2})$ and input $X \sim \text{Bernoulli}(\frac{1}{2})$.

$$Z = \begin{cases} \log 2(1-\delta) & \text{w.p. } 1-\delta, \\ \log 2\delta & \text{w.p. } \delta, \end{cases}$$

The CGF of $\tilde{S}_n$ is $K_{\tilde{S}_n}(s) = n \log\left(\delta + (1-\delta) e^{\log \frac{1-\delta}{\delta} s}\right)$. Solving $K'_{\tilde{S}_n}(s) = k - n\log 2\delta$ gives a closed form expression for the saddlepoint:

$$\hat{s} = \frac{\log\left(\frac{(k/n - \log 2\delta)\delta}{(\log(2(1-\delta)) - k/n)(1-\delta)}\right)}{\log\left(\frac{1-\delta}{\delta}\right)}.$$

Binary Erasure Channel (BEC): Similar derivations.

Takeaway:

Saddlepoint $\hat{s}$ is in closed form with respect to γ and n for common channels.