Nonasymptotic Oblivious Relaying and Variable-Length Noisy Lossy Source Coding

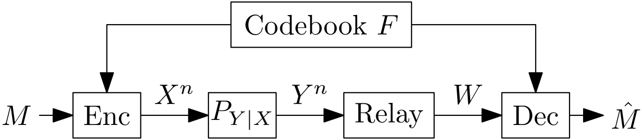
Yanxiao Liu*, Sepehr Heidari Advary† and Cheuk Ting Li*

* The Chinese University of Hong Kong
† Sharif University of Technology

2025 IEEE International Symposium on Information Theory

June 2025

Oblivious Relay Channel (Sanderovich et al., 2008; Simeone et al., 2011)

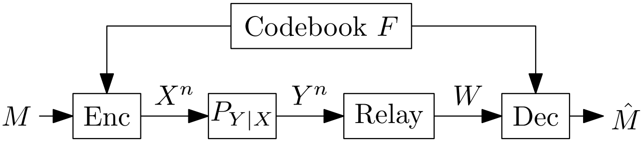


In the oblivious relay channel / information bottleneck channel:

- An encoder encodes a message M into $X^n = (X_1, \ldots, X_n)$
- X^n is sent through a memoryless channel $P_{Y|X}$
- A relay observes Y^n, and relays it via rate-limited description W
  - Relay is **oblivious**—does not know the codebook
  - Relay can only perform lossy compression on Y^n without decoding it
- Decoder observes W and attempts to decode M

Application: cloud radio access networks (Aguerri et al., 2019)

- Base stations are connected to a cloud-computing central processor via error-free rate-limited fronthaul links



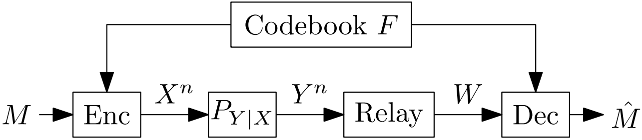
Asymptotic Capacity (Sanderovich et al., 2008)

As $n \to \infty$, minimum description rate needed to support a message transmission rate C is given by the **information bottleneck**

$$\mathrm{IB}(\mathsf{C}) := \min_{P_{U|Y}:\, I(X;U) \geq \mathsf{C}} I(Y;U), \qquad X \leftrightarrow Y \leftrightarrow U$$

Intuition:

- If we want to transmit M at rate C, we need $I(X^n;W) \geq n\mathsf{C}$
- If description rate is R_W, then $I(Y^n;W) \leq n\mathsf{R}_W$
- Minimum R_W is $\approx \min_{P_{W|Y^n}:\, n^{-1}I(X^n;W) \geq \mathsf{C}} \; n^{-1}I(Y^n;W)$

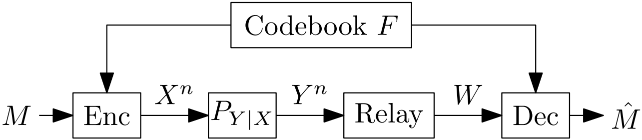


Asymptotic Capacity (Sanderovich et al., 2008)

As $n \to \infty$, minimum description rate needed to support a message transmission rate C is given by the **information bottleneck**

$$\mathrm{IB}(\mathsf{C}) := \min_{P_{U|Y}:\, I(X;U) \geq \mathsf{C}} I(Y;U), \qquad X \leftrightarrow Y \leftrightarrow U$$

- The intuition only holds for $n \to \infty$
- In practice, we always have a finite blocklength n
- Also see Wu and Joudeh (2024) for an error exponent analysis
- **Our contribution:** Nonasymptotic achievability results



- For finite n, it matters whether W is **fixed-length** or **variable-length**

Theorem (Nonasymptotic Achievability for Fixed-Length Description)

For error probability ϵ, a fixed-length description rate

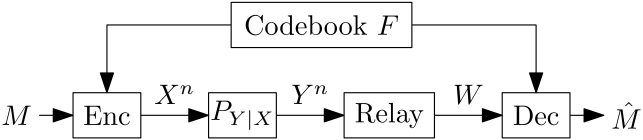
$$\mathrm{IB}(\mathrm{C}) + \sqrt{\frac{1}{n}\mathrm{VIB}(\mathrm{C})}\,Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

suffices, where

$$\mathrm{VIB}(\mathrm{C}) := \mathrm{Var}\left[\iota_{Y;U}(Y;U) - \lambda^* \iota_{X;U}(X;U)\right]$$

is a "second-order information bottleneck", $\iota_{Y;U}(y;u) := \log\left(P_{U|Y}(u|y)/P_U(u)\right)$ is computed using the optimal $P_{U|Y}$ in $\mathrm{IB}(\mathrm{C})$, and $\lambda^* := \frac{\mathrm{d}}{\mathrm{dC}}\mathrm{IB}(\mathrm{C})$

Nonasymptotic Oblivious Relay Channel



- For finite n, it matters whether W is **fixed-length** or **variable-length**

Theorem (Nonasymptotic Achievability for Variable-Length Description)

For error probability ϵ, for a variable-length $W \in \{0,1\}^*$ in a prefix code, a description rate

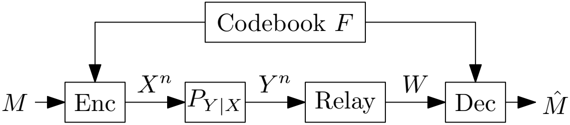
$$(1-\epsilon)\left(\mathrm{IB}(\mathrm{C}) + \sqrt{\frac{\ln n}{n}\mathrm{CVIB}(\mathrm{C})}\right) + O\left(\frac{1}{\sqrt{n}}\right)$$

suffices, where

$$\mathrm{CVIB}(\mathrm{C}) := \mathbb{E}\left[\mathrm{Var}\left[\lambda^* \iota_{X;U}(X;U) \,\middle|\, Y, U\right]\right]$$

is a "conditional second-order IB"

Nonasymptotic Oblivious Relay Channel



- **Fixed-length**:

$$\mathrm{IB}(\mathrm{C}) + \sqrt{\frac{1}{n}\mathrm{VIB}(\mathrm{C})}Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

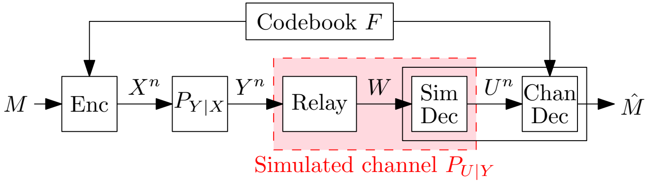
  where $\mathrm{VIB}(\mathrm{C}) := \mathrm{Var}\left[\iota_{Y;U}(Y;U) - \lambda^* \iota_{X;U}(X;U)\right]$

- **Variable-length**:

$$(1-\epsilon)\left(\mathrm{IB}(\mathrm{C}) + \sqrt{\frac{\ln n}{n}\mathrm{CVIB}(\mathrm{C})}\right) + O\left(\frac{1}{\sqrt{n}}\right)$$

  where $\mathrm{CVIB}(\mathrm{C}) := \mathbb{E}\left[\mathrm{Var}\left[\lambda^* \iota_{X;U}(X;U) \,\middle|\, Y, U\right]\right]$

- The two cases have vastly different behaviors
  - This phenomenon was also observed in lossy source coding (Kostina et al., 2015)

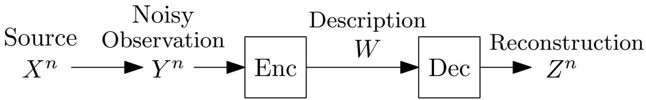


- Asymptotic rate: $\mathrm{IB}(\mathrm{C}) = \min_{P_{U|Y}: I(X;U) \geq \mathrm{C}} I(Y; U)$
- Idea: Use W to simulate $P_{U|Y}$ via **channel simulation**
  - Decoder treat $X^n \to U^n$ as a channel $P_{U|X} = P_{Y|X}P_{U|Y}$, and use ordinary channel decoder
- One-shot variable-length channel simulation via strong functional representation lemma (Li and El Gamal, 2018; Li, 2024):

$$\mathbb{E}[|W|] \leq nI(Y; U) + \log(nI(Y; U) + 2) + 3$$

  - Good enough to give asymptotic rate, but not enough for nonasymptotic
- Idea 2: Use **nonasymptotic noisy lossy source coding** (Kostina and Verdú, 2016) together with **Poisson matching lemma** (Li and Anantharam, 2021)

Noisy Lossy Source Coding (Dobrushin and Tsybakov, 1962)



- 2-discrete memoryless source X^n, Y^n
- Encoder observes Y^n, sends description W to decoder
- Decoder recovers Z^n, wants $\mathbb{P}(n^{-1}\sum_i d(X_i, Z_i) > \mathrm{D}) \le \epsilon$
- Optimal asymptotic rate is (Dobrushin and Tsybakov, 1962)

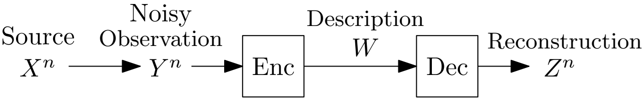
$$R(\mathrm{D}) := \min_{P_{Z|Y}:\, \mathbb{E}[d(X,Z)] \le \mathrm{D}} I(Y;Z)$$

- For finite n, we again have two cases for W
- **Fixed-length**: Optimal length is (Kostina and Verdú, 2016)

$$nR(\mathrm{D}) + \sqrt{n\tilde{\mathrm{V}}(\mathrm{D})}Q^{-1}(\epsilon) + O(\log n)$$

where $\tilde{\mathrm{V}}(\mathrm{D}) := \mathrm{Var}[\iota_{Y;Z^*}(Y;Z^*) + \lambda^* d(X, Z^*)]$, $P_{Z^*|Y}$ attains the minimum in $R(\mathrm{D})$, and $\lambda^* := -R'(\mathrm{D})$

Noisy Lossy Source Coding (Dobrushin and Tsybakov, 1962)



- Optimal asymptotic rate is $R(\mathrm{D}) := \min_{P_{Z|Y}:\mathbb{E}[d(X,Z)]\le \mathrm{D}} I(Y;Z)$
- **Fixed-length**: Optimal length is (Kostina and Verdú, 2016)

$$nR(\mathrm{D}) + \sqrt{n\tilde{\mathrm{V}}(\mathrm{D})}Q^{-1}(\epsilon) + O(\log n)$$

where $\tilde{\mathrm{V}}(\mathrm{D}) := \mathrm{Var}[\iota_{Y;Z^*}(Y;Z^*) + \lambda^* d(X,Z^*)]$, $\lambda^* := -R'(\mathrm{D})$

- **Variable-length**: We proved:

Theorem (Noisy Lossy Source Coding w/ Variable-Length Description)

For variable-length prefix-free description, we can have an expected length

$$(1-\epsilon)\left(nR(\mathrm{D}) + \sqrt{(n\ln n)\widetilde{\mathrm{CV}}(\mathrm{D})}\right) + O(\sqrt{n}),$$

where $\widetilde{\mathrm{CV}}(\mathrm{D}) := (\lambda^*)^2\mathbb{E}[\mathrm{Var}[d(X,Z^*)\,|\,Y,Z^*]]$

One-shot Noisy Lossy Source Coding

- We first show the following one-shot result ($n = 1$), via the technique in (Kostina and Verdú, 2016) and the Poisson functional representation (PFR) (Li and El Gamal, 2018; Li and Anantharam, 2021)

Theorem (One-shot Lossy Source Coding w/ Variable-Length Description)

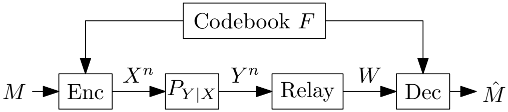
For any $P_{\bar{Z}}$, $\epsilon' > 0$, and function $\beta : \mathcal{Y} \to [0, 1]$, there is a variable-length code with $P_e \leq \mathbb{E}[\beta(Y)] + \epsilon'$ and

$$\mathbb{E}[|W|] \leq \ell\left(\mathbb{E}\left[(1 - \beta(Y))\psi_{\bar{Z}}(Y, \mathrm{D}, \epsilon')\right]\right),$$

$$\psi_{\bar{Z}}(y, \mathrm{D}, t) = \inf_{P_Z : \mathbb{P}(d(X,Z) > \mathrm{D} | Z, Y = y) \leq t \text{ a.s.}} D(P_Z \| P_{\bar{Z}}), \; \ell(t) = t + \log(t + 2) + 4$$

- PFR on the channel $P_{\hat{Z}|Y}$, where conditional on $Y = y$, $\hat{Z}$ has the same distribution as $\bar{Z} \sim P_{\bar{Z}}$ conditional on $\phi(y, \bar{Z}, \mathrm{D}) \leq \epsilon'$, $\phi(y, z, \mathrm{D}) := \mathbb{P}(d(X, z) > \mathrm{D} | Y = y)$
- Encoder produces index of PFR with probability $1 - \beta(Y)$, or index 1 with probability $\beta(Y)$, then encodes index into W

Nonasymptotic Variable-Length Oblivious Relay Channel



Theorem (Nonasymptotic Achievability for Variable-Length Description)

$$\text{Achievable rate: } (1-\epsilon)\left(\mathrm{IB}(\mathsf{C}) + \sqrt{\frac{\ln n}{n}\mathrm{CVIB}(\mathsf{C})}\right) + O\left(\frac{1}{\sqrt{n}}\right),$$

$$\mathrm{CVIB}(\mathsf{C}) := \mathbb{E}\left[\mathrm{Var}\left[\lambda^* \iota_{X;U}(X;U) \,\middle|\, Y, U\right]\right]$$

Idea: Define distortion function, where small distortion $\Rightarrow$ small error prob.

- Define $d(x,u) = -\iota_{X;U}(x;u)$, and consider rate-distortion function $R(\mathsf{D}) = \min_{P_{\tilde{U}|Y}:\mathbb{E}[d(X,\tilde{U})]\le \mathsf{D}} I(Y;\tilde{U})$ of noisy lossy source coding
- $R(-\mathsf{C}) = \mathrm{IB}(\mathsf{C})$
- Noisy lossy source coding: Decoder recovers $\hat{U}^n$ where $\sum_i \iota_{X;U}(X_i;\hat{U}_i) \ge n\mathsf{C}$ with high prob.

Nonasymptotic Variable-Length Oblivious Relay Channel

Theorem (Nonasymptotic Achievability for Variable-Length Description)

$$\text{Achievable rate: } (1-\epsilon)\left(\mathrm{IB(C)}+\sqrt{\frac{\ln n}{n}\mathrm{CVIB(C)}}\right)+O\left(\frac{1}{\sqrt{n}}\right),$$

$$\mathrm{CVIB(C)} := \mathbb{E}\left[\mathrm{Var}\left[\lambda^* \iota_{X;U}(X;U) \,\middle|\, Y, U\right]\right]$$

- Decoder recovers $\hat{U}^n$ where $\sum_i \iota_{X;U}(X_i;\hat{U}_i) \geq n\mathrm{C}$ with high prob.
- $\hat{U}^n$ has "enough information" about X^n for decoding $n\mathrm{C}$ message bits
- Use Poisson functional representation: $\bar{X}_1^n, \bar{X}_2^n, \ldots \overset{iid}{\sim} P_{X^n}$, Poisson process $0 < T_1 < T_2, \ldots$
- Encoder sends $X^n = \bar{X}_M^n$
- Decoder recovers $\hat{U}^n$ via noisy lossy source coding and recovers

$$\hat{M} := \mathrm{argmin}_k T_k / \left(P_{X^n|U^n}(\bar{X}_k^n|\hat{U}^n)/P_{X^n}(\bar{X}_k^n)\right)$$

- Poisson matching lemma: Decodes correctly with high probability

Conclusion

We proved:

- Nonasymptotic achievability for oblivious relay channel
  - Fixed and variable-length description
- Nonasymptotic achievability for variable-length noisy lossy source coding

Future directions:

- Second order converses—are these results tight?
- Error exponent for variable-length description

Acknowledgement

This work was partially supported by two grants from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No.s: CUHK 24205621 (ECS), CUHK 14209823 (GRF)].

References

Aguerri, I. E., Zaidi, A., Caire, G., and Shamai, S. (2019). On the capacity of cloud radio access networks with oblivious relaying. *IEEE Transactions on Information Theory*, 65(7):4575–4596.

Dobrushin, R. and Tsybakov, B. (1962). Information transmission with additional noise. *IRE Transactions on Information Theory*, 8(5):293–304.

Kostina, V., Polyanskiy, Y., and Verdú, S. (2015). Variable-length compression allowing errors. *IEEE Transactions on Information Theory*, 61(8):4316–4330.

Kostina, V. and Verdú, S. (2016). Nonasymptotic noisy lossy source coding. *IEEE Transactions on Information Theory*, 62(11):6111–6123.

Li, C. T. (2024). Channel simulation: Theory and applications to lossy compression and differential privacy. *Foundations and Trends® in Communications and Information Theory*, 21(6):847–1106.

Li, C. T. and Anantharam, V. (2021). A unified framework for one-shot achievability via the Poisson matching lemma. *IEEE Transactions on Information Theory*, 67(5):2624–2651.

Li, C. T. and El Gamal, A. (2018). Strong functional representation lemma and applications to coding theorems. *IEEE Transactions on Information Theory*, 64(11):6967–6978.

Sanderovich, A., Shamai, S., Steinberg, Y., and Kramer, G. (2008). Communication via decentralized processing. *IEEE Transactions on Information Theory*, 54(7):3008–3023.

Simeone, O., Erkip, E., and Shamai, S. (2011). On codebook information for interference relay channels with out-of-band relaying. *IEEE Transactions on Information Theory*, 57(5):2880–2888.

Wu, H. and Joudeh, H. (2024). An achievable error exponent for the information bottleneck channel. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1297–1302. IEEE.