Pointwise Redundancy in One-Shot Lossy Compression via Poisson Functional Representation

Conference Paper

Author(s):
Li, Cheuk Ting

Publication date:
2024-03-06

Permanent link:
https://doi.org/10.3929/ethz-b-000664571

International Zurich Seminar on Information and Communication (IZS), March 6 – 8, 2024

Pointwise Redundancy in One-Shot Lossy Compression via Poisson Functional Representation

Cheuk Ting Li
The Chinese University of Hong Kong
Hong Kong SAR, China
email: ctli@ie.cuhk.edu.hk

***Abstract*—We present a construction of one-shot variable-length lossy source coding schemes using the Poisson functional representation, and give bounds on its pointwise redundancy. This allows us to describe the distribution of the encoding length in a precise manner.**

I. Introduction

Variable-length lossy source coding has been considered, for example, in D-semifaithful codes [1], [2] where the distortion must be bounded almost surely. The redundancy of D-semifaithful codes, i.e., the difference between the encoding length and the rate distortion function, has been studied in [3]–[6].

For one-shot variable-length lossy source coding with the expected distortion constraint $\mathbb{E}[d(X,Y)] \le D$,[1] it was proved in [7] that there is a prefix-free code with expected length $\le R(D) + \log(R(D)+1) + 6$, showing that the optimal one-shot expected length is always within a logarithmic gap from the rate-distortion function $R(D)$. The proof utilizes the Poisson functional representation [7], [8], where the codebook is constructed as a Poisson process. Also see [9]–[11] for related results.

In this work, we utilize the Poisson functional representation to construct one-shot variable-length lossy source coding schemes, and give bounds on their pointwise redundancy. This allows us to describe the distribution of the encoding length in a more precise manner, compared to only bounding its expectation. The proofs and details of the results mentioned in this abstract, and the generalization to the lossy Gray-Wyner system [12], can be found in the preprint [13].

II. Main Results

A one-shot variable-length lossy compression scheme for the source $X \in \mathcal{X}$, $X \sim P_X$ with reconstruction space $\mathcal{Y}$ is a pair $(P_{M|X}, g)$, where $P_{M|X}$ is a stochastic encoder (a conditional distribution from $\mathcal{X}$ to $\{0,1\}^*$, where $\{0,1\}^*$ is the set of bit sequences of any length), and $g : \{0,1\}^* \to \mathcal{Y}$ is a decoding function. The encoder observes $X \sim P_X$ and outputs the description $M|X \sim P_{M|X}$. The decoder observes M and outputs the reconstruction $\tilde{Y} = g(M)$. We can choose whether to impose the prefix-free condition on M or not. We may impose an expected distortion constraint $\mathbb{E}[d(X,\tilde{Y})] \le D$, where $d : \mathcal{X} \times \mathcal{Y} \to [0,\infty)$ is a distortion function.

We can also replace the variable-length description M by a positive integer K, and assume that the encoder produces a positive integer description. Note that we can convert K into a variable-length description with $\lfloor \log K \rfloor$ bits without the prefix-free condition [14], or $\le \log K + 2\log(\log K + 1) + 1$ bits with the prefix-free condition using the Elias delta code [15].

The following theorem can be proved using the Poisson functional representation construction similar to [7, Theorem 2], with an analysis using techniques in [8]. Refer to [13] for the proof.

Theorem 1: Fix any P_X, $P_{Y|X}$ and Q_Y satisfying $P_{Y|X}(\cdot|x) \ll Q_Y$ for P_X-almost all x's. Fix any collection of functions $\psi_i : \mathcal{X} \times \mathcal{Y} \times \mathbb{Z}_{>0} \to \mathbb{R}$ that are nondecreasing in the third argument for $i = 1, \ldots, \ell$. Then there exists a lossy compression scheme with description $K \in \mathbb{Z}_{>0}$ and reconstruction $\tilde{Y}$ such that

$$\mathbb{E}\big[\psi_i(X,\tilde{Y},K)\big] \le \mathbb{E}\big[\psi_i(X,Y,\ell J)\big]$$

for $i = 1, \ldots, \ell$, where $(X,Y) \sim P_X P_{Y|X}$, and $J \in \mathbb{Z}_{>0}$ is distributed as

$$J|(X,Y) \sim \mathrm{Geom}\bigg(\bigg(\frac{\mathrm{d}P_{Y|X}(\cdot|X)}{\mathrm{d}Q_Y}(Y) + 1\bigg)^{-1}\bigg).$$

This theorem is quite general. For example, to bound the expected distortion, take $\psi_i(x,y,k) = d(x,y)$. To bound the excess distortion probability, take $\psi_i(x,y,k) = \mathbf{1}\{d(x,y) > D\}$. To bound the probability that K cannot be encoded into n bits (for a fixed-length code), take $\psi_i(x,y,k) = \mathbf{1}\{k > 2^n\}$. To bound the expected length with (resp. without) the prefix-free condition, we may take $\psi_i(x,y,k) = \log k$ (resp. $\psi_i(x,y,k) = \log k + 2\log(\log k + 1) + 1$).

We can also use Theorem 1 to bound the pointwise redundancy. We consider three different notions of pointwise redundancy: **Pointwise rate redundancy (PRR)**, studied in [5], [16], is given by

$$|M| - R(D),$$

i.e., the difference between the length $|M|$ of the description M and the rate-distortion function $R(D)$ where $D = \mathbb{E}[d(X,\tilde{Y})]$.

This work was partially supported by an ECS grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No.: CUHK 24205621].

[1]Note that the probability of excess distortion $\mathbb{P}(d(X,Y) > D) = \mathbb{E}[\mathbf{1}\{d(X,Y) > D\}]$ can also be written as an expected distortion.

Pointwise source-wise redundancy (PSR), studied in [5], is given by

$$|M| - \jmath(X, D),$$

where $\jmath(x, D)$ is the *d-tilted information* [5], [17], [18] $\jmath(x, D) := -\log \mathbb{E}[2^{-\lambda^*(d(x,Y^*)-D)}]$, where $Y^* \sim P_Y$ follows the Y-marginal of $P_X P_{Y|X}$ where $P_{Y|X}$ is the conditional distribution that attains the minimum in $R(D)$ (assume unique minimizer), and $\lambda^* := -R'(D)$. **Pointwise source-distortion-wise redundancy (PSDR)** is defined as

$$|M| - \jmath(X, D, d(X, \tilde{Y})),$$

where we write $\jmath(x, D, \delta) := -\log \mathbb{E}[2^{-\lambda^*(d(x,Y^*)-\delta)}] = \jmath(x, D) - \lambda^*(\delta - D)$, which can be interpreted as the amount of information needed to convey x within a distortion δ when the overall expected distortion is D. The expectations of these three redundancies must be nonnegative for prefix-free codes, but might be negative if we do not impose the prefix-free condition. We first state a corollary of Theorem 1 that can bound any of the three pointwise redundancies for the case without the prefix-free condition.

Corollary 2: Fix any P_X, $P_{Y|X}$, distortion function $d : \mathcal{X} \times \mathcal{Y} \to [0, \infty)$, function $\eta : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ and $\gamma \in \mathbb{R}$. Then there exists a lossy compression scheme without prefix-free condition such that $\mathbb{E}[d(X, \tilde{Y})] \le \mathbb{E}[d(X, Y)]$, and

$$\mathbb{P}\left(|M| - \eta(X, \tilde{Y}) \ge \gamma\right) \le \mathbb{E}\left[\min\left\{2^{-\eta(X,Y)-\gamma+1}(2^{\iota_{X;Y}(X;Y)} + 1),\, 1\right\}\right],$$

where $(X, Y) \sim P_X P_{Y|X}$.

The result for PSDR is especially simple.

Corollary 3: For $D > 0$, under the regularity conditions in [18],[2] there exists a lossy compression scheme without prefix-free condition, with $\mathbb{E}[d(X, \tilde{Y})] \le D$, and with PSDR satisfying

$$\mathbb{P}\left(|M| - \jmath(X, D, d(X, \tilde{Y})) \ge \gamma\right) \le 2^{-\gamma+2}$$

for every $\gamma \in \mathbb{R}$.

The results for prefix-free codes are slightly more complicated.

Corollary 4: Fix any P_X, $P_{Y|X}$, distortion function $d : \mathcal{X} \times \mathcal{Y} \to [0, \infty)$, function $\eta : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, and $\gamma \in \mathbb{R}$. Then there exists a prefix-free lossy compression scheme such that $\mathbb{E}[d(X, \tilde{Y})] \le \mathbb{E}[d(X, Y)]$, and

$$\mathbb{P}\left(|M| - \eta(X, \tilde{Y}) \ge \gamma\right) \le \mathbb{E}\Big[\min\big\{2^{-\eta(X,Y)-\gamma+2}([\eta(X, Y) + \gamma]_+ + 1)^2 \cdot (2^{\iota_{X;Y}(X;Y)} + 1),\, 1\big\}\Big],$$

where $(X, Y) \sim P_X P_{Y|X}$.

[2]The regularity conditions in [18] are: $R(\delta)$ is finite for some δ, there exists a finite set $\mathcal{E} \subseteq \mathcal{Y}$ such that $\mathbb{E}[\min_{y \in \mathcal{E}} d(X, y)] < \infty$, and the minimum in $R(D)$ is achieved by a unique $P_{Y|X}$.

Corollary 5: For $D > 0$, $\gamma \in \mathbb{R}$, under the regularity conditions in [18] (see Corollary 3), there exists a prefix-free lossy compression scheme with $\mathbb{E}[d(X, \tilde{Y})] \le D$, and with PSDR satisfying

$$\mathbb{P}\left(|M| - \jmath(X, D, d(X, \tilde{Y})) \ge \gamma\right) \le 2^{-\gamma+3}\mathbb{E}\big[([\iota_{X;Y}(X; Y) + \gamma]_+ + 1)^2\big].$$

References

[1] D. S. Ornstein and P. C. Shields, "Universal almost sure data compression," *The Annals of Probability*, pp. 441–452, 1990.

[2] B. Yu and T. P. Speed, "A rate of convergence result for a universal d-semifaithful code," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 813–820, 1993.

[3] J. C. Kieffer, "Sample converses in source coding theory," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 263–268, 1991.

[4] Z. Zhang, E.-H. Yang, and V. K. Wei, "The redundancy of source coding with a fidelity criterion. 1. known statistics," *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 71–91, 1997.

[5] I. Kontoyiannis, "Pointwise redundancy in lossy data compression and universal lossy data compression," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 136–152, 2000.

[6] A. Dembo and I. Kontoyiannis, "Critical behavior in lossy source coding," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1230–1236, 2001.

[7] C. T. Li and A. El Gamal, "Strong functional representation lemma and applications to coding theorems," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 6967–6978, Nov 2018.

[8] C. T. Li and V. Anantharam, "A unified framework for one-shot achievability via the Poisson matching lemma," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2624–2651, 2021.

[9] E. C. Posner and E. R. Rodemich, "Epsilon entropy and data compression," *The Annals of Mathematical Statistics*, pp. 2079–2125, 1971.

[10] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 438–449, Jan 2010.

[11] M. Braverman and A. Garg, "Public vs private coin in bounded-round information," in *International Colloquium on Automata, Languages, and Programming*. Springer, 2014, pp. 502–513.

[12] R. Gray and A. Wyner, "Source coding for a simple network," *Bell System Technical Journal*, vol. 53, no. 9, pp. 1681–1721, 1974.

[13] C. T. Li, "Pointwise redundancy in one-shot lossy compression via Poisson functional representation," *arXiv preprint*, 2024.

[14] W. Szpankowski and S. Verdú, "Minimum expected length of fixed-to-variable lossless compression without prefix constraints," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4017–4025, 2011.

[15] P. Elias, "Universal codeword sets and representations of the integers," *IEEE transactions on information theory*, vol. 21, no. 2, pp. 194–203, 1975.

[16] B. Oğuz and V. Anantharam, "Pointwise lossy source coding theorem for sources with memory," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 363–367.

[17] I. Csiszár, "On an extremum problem of information theory," *Studia Scientiarum Mathematicarum Hungarica*, vol. 9, no. 1, pp. 57–71, 1974.

[18] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3309–3338, 2012.