Full text available at: http://dx.doi.org/10.1561/0100000141

Channel Simulation: Theory and Applications to Lossy Compression and Differential Privacy

Other titles in Foundations and Trends® in Communications and Information Theory

Maximizing Entropy with an Expectation Constraint and One-Parameter Exponential Families of Distributions: A Reexamination
David L. Neuhoff
ISBN: 978-1-63828-480-2

Codes for Adversaries: Between Worst-Case and Average-Case Jamming
Bikash Kumar Dey, Sidharth Jaggi, Michael Langberg, Anand D. Sarwate and Yihan Zhang
ISBN: 978-1-63828-460-4

Universal Features for High-Dimensional Learning and Inference
Shao-Lun Huang, Anuran Makur, Gregory W. Wornell and Lizhong Zheng
ISBN: 978-1-63828-176-4

Ultra-Reliable Low-Latency Communications: Foundations, Enablers, System Design, and Evolution Towards 6G
Nurul Huda Mahmood, Italo Atzeni, Eduard Axel Jorswieck and Onel Luis Alcaraz López
ISBN: 978-1-63828-180-1

Probabilistic Amplitude Shaping
Georg Böcherer
ISBN: 978-1-63828-178-8

Finite Blocklength Lossy Source Coding for Discrete Memoryless Sources
Lin Zhou and Mehul Motani
ISBN: 978-1-63828-182-5

Channel Simulation: Theory and Applications to Lossy Compression and Differential Privacy

Cheuk Ting Li
The Chinese University of Hong Kong
ctli@ie.cuhk.edu.hk

the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

C. T. Li. *Channel Simulation: Theory and Applications to Lossy Compression and Differential Privacy*. Foundations and Trends® in Communications and Information Theory, vol. 21, no. 6, pp. 847–1106, 2024.

ISBN: 978-1-63828-487-1

Foundations and Trends® in Communications and Information Theory
Volume 21, Issue 6, 2024

Editorial Scope

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2024, Volume 21, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Preface 2

Notations 3

1 Introduction and Motivations 7

1.1 Overview 7
1.2 History 13
1.3 Dithered Quantization 14
1.4 Machine Learning 16
1.5 Lossy Compression 19
1.6 Privacy 22
1.7 Distribution Preserving Quantization and Rate-Distortion-Perception Tradeoff 26
1.8 Simulating Quantum Measurements via Classical Communication 28
1.9 Communication Complexity and Message Compression . 31
1.10 Coordination Problems 34
1.11 Other Applications 36
1.12 Preliminaries—Prefix-free Codes 36

2 The Channel Simulation Setting **39**
2.1 Definition and Notations 39
2.2 Overview of Various Channel Simulation Approaches . . 46
2.3 Why Unlimited Common Randomness? Why Prefix-free? Why Expected Length? 50

3 One-shot Channel Simulation with Unlimited Common Randomness **56**
3.1 Functional Representation Lemma 58
3.2 Rejection Sampling 61
3.3 Exponential and Poisson Functional Representation . . . 75
3.4 Likelihood Encoder and Minimal Random Coding 91
3.5 Discussions on Sampling Schemes 96
3.6 Subtractively Dithered Quantization and Universal Quantization 103
3.7 Exact Fixed-Length Channel Simulation 119

4 One-shot Channel Simulation without Common Randomness **121**
4.1 Discrete Channels 123
4.2 Continuous Channels 125

5 Asymptotic Channel Simulation with Unlimited Common Randomness **133**
5.1 Exact Variable-Length Channel Simulation 133
5.2 Method of Types 135
5.3 Total Variation Distance 137
5.4 Approximate Fixed-Length Channel Simulation 145
5.5 Soft Covering Lemma 150
5.6 Likelihood Encoder for Asymptotic Coding 152

6 Asymptotic Channel Simulation without Common Randomness **155**
6.1 Approximate Case—Wyner's Common Information 155
6.2 Exact Case—Exact Common Information Rate 158

7 Asymptotic Channel Simulation with Limited Common Randomness 163
7.1 Approximate Channel Simulation 163
7.2 Exact Channel Simulation 167

8 One-shot Bounds for Fixed-Length Channel Simulation 169
8.1 One-shot Soft Covering Lemma 171
8.2 One-shot Fixed-length Channel Simulation 174

9 Source and Channel Simulation with Limited Local Randomness 180
9.1 Source Simulation 180
9.2 Distributed Source Simulation 187
9.3 Local Channel Simulation 193
9.4 Channel Simulation with Limited Common and Local Randomness 204

10 Other Settings 209
10.1 Simulating a Channel with Feedback 209
10.2 Simulating a Channel using Another Channel 211
10.3 Interactive Channel Simulation 215
10.4 Secure Channel Simulation 217
10.5 Channel Simulation over Networks 218
10.6 Single-Input Multiple-Output Channel Simulation 219

11 Conclusions and Future Directions 221

Acknowledgements 225

Appendices 226

References 233

Channel Simulation: Theory and Applications to Lossy Compression and Differential Privacy

Cheuk Ting Li

Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, China; ctli@ie.cuhk.edu.hk

ABSTRACT

One-shot channel simulation (or channel synthesis) has seen increasing applications in lossy compression, differential privacy and machine learning. In this setting, an encoder observes a source X, and transmits a description to a decoder, so as to allow it to produce an output Y with a desired conditional distribution $P_{Y|X}$. In other words, the encoder and the decoder are simulating the noisy channel $P_{Y|X}$ using noiseless communication. This can also be seen as a lossy compression scheme with a stronger guarantee on the joint distribution of X and Y. This monograph gives an overview of the theory and applications of the channel simulation problem. We will present a unifying review of various one-shot and asymptotic channel simulation techniques that have been proposed in different areas, namely dithered quantization, rejection sampling, minimal random coding, likelihood encoder, soft covering, Poisson functional representation, and dyadic decomposition.

Cheuk Ting Li (2024), "Channel Simulation: Theory and Applications to Lossy Compression and Differential Privacy", Foundations and Trends® in Communications and Information Theory: Vol. 21, No. 6, pp 847–1106. DOI: 10.1561/0100000141.

Preface

In this monograph, we give an overview of the theoretical results on channel simulation and related settings, as well as their applications in lossy compression, differential privacy and machine learning. We collect various channel simulation schemes appearing in different fields of research. Many of them are not referred to as “channel simulation” in their respective fields. Nevertheless, they fit within the same setting of simulating a noisy channel through communications, and can therefore be analyzed and compared under a unified framework. Our goal is to gather these channel simulation techniques, and present them as a common toolbox that different lines of research can utilize.

Although this monograph is intended to be accessible to researchers outside of information theory, familiarity with basic notions such as entropy, mutual information and channel coding is necessary. Readers may consult textbooks such as [58, Chapters 1-10]; [176, Chapters 1-11]; [280, Chapters 1-11]; or [61, Chapters 1-7].

In Section 1, we will give an intuitive description of the channel simulation setting, and present several motivations for this setting. Readers using this monograph as a reference book may jump directly to the overview of various channel simulation schemes in Section 2.2, and the comparison in Table 2.1.

Appendices

A

Zipf Distribution

The *Zipf distribution* [215] (also known as *zeta distribution*) with parameter $s > 1$ is a distribution over $\mathbb{N}^+$ with probability mass function

$$\mathrm{Zipf}(k; s) := \frac{k^{-s}}{\zeta(s)},$$

where $\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$ is the Riemann zeta function. It is the maximum entropy distribution for $K \in \mathbb{N}^+$ when $\mathbb{E}[\log_2 K]$ is fixed. We can use the Zipf distribution to show the following bound (e.g., see [155]).

Proposition 75. *For random variable $K \in \mathbb{N}^+$ following the distribution P_K, its cross entropy with $\mathrm{Zipf}(s)$ is bounded by*

$$H(P_K, \mathrm{Zipf}(s)) \leq s\mathbb{E}[\log_2 K] + \log_2 \frac{s}{s-1}. \tag{A.1}$$

Therefore, if $\mathbb{E}[\log_2 K] \leq \ell$, letting $s = 1 + 1/\ell$, we have

$$H(K) \leq H(P_K, \mathrm{Zipf}(s)) \leq \ell + \log_2(\ell + 1) + 1.$$

Proof. We have

$$H(P_K, \mathrm{Zipf}(s)) = \sum_{k=1}^{\infty} P_K(k) \log_2 \frac{\zeta(s)}{k^{-s}}$$

$$= s\mathbb{E}[\log_2 K] + \log_2 \zeta(s),$$

where

$$\zeta(s) \le 1 + \int_1^\infty \kappa^{-s} \mathrm{d}\kappa = \frac{s}{s-1}. \tag{A.2}$$

The result follows. □

Proposition 75 suggests that, if we know that $\mathbb{E}[\log_2 K] \le \ell$, then we can use the Shannon code [228] designed for the distribution Zipf(s) where $s = 1 + 1/\ell$ to encode K, to obtain a codeword with expected length upper-bounded by

$$H(P_K, \mathrm{Zipf}(s)) + 1 \le \ell + \log_2(\ell + 1) + 2 \text{ bits}.$$

Refer to Section 1.12. The downside is that we need to know ℓ when we construct the code, and the Shannon code over an infinite alphabet can be hard to construct.

In contrast, if we do not know the bound $\mathbb{E}[\log_2 K] \le \ell$ when we design the code, we can still use the Elias delta code [75] to encode K, which will result in a codeword length upper-bounded by

$$\ell + 2\log_2(\ell + 1) + 1 \text{ bits}$$

if $\mathbb{E}[\log_2 K] \le \ell$. While it is possible to improve this bound to $\ell + (1 + \epsilon)\log_2(\ell + 1) + O(1)$, for example, by using the Elias omega code [75], it is impossible to design a prefix-free code over $\mathbb{N}^+$ that achieves an expected length upper-bounded by $\ell + \log_2(\ell + 1) + O(1)$ for every ℓ and random variable K with $\mathbb{E}[\log_2 K] \le \ell$.[1] Therefore, although using a "universal" code such as the Elias delta code has the advantage that we do not need to know the bound $\mathbb{E}[\log_2 K] \le \ell$ beforehand, it comes with a small penalty on the expected length.

Practically, if we are given the bound $\mathbb{E}[\log_2 K] \le \ell$, then there are several options for the encoding of $K \in \mathbb{N}^+$:

- Shannon code [228] for the distribution Zipf$(1+1/\ell)$, or any prefix-free code $f : \mathbb{N}^+ \to \{0,1\}^*$ with $|f(k)| \le \lceil \mathrm{Zipf}(k; 1 + 1/\ell) \rceil$ for $k \in \mathbb{N}^+$. The expected length is upper-bounded by $\ell + \log_2(\ell+1) + 2$. Nevertheless, it can be hard to construct.

[1] This is because $\sum_{k=1}^\infty 2^{-\log_2 k - \log_2(\log_2 k + 1) - c} = \sum_{k=1}^\infty \frac{1}{2^c k(\log_2 k + 1)} = \infty$, violating Kraft's inequality [142].

- A code over positive integers with efficient encoding and decoding algorithms such as the Elias delta code [75], with a slight penalty on the expected length. The advantage is that the code does not depend on ℓ.

- Use a "hybrid" approach: first construct the Shannon code $f_{\mathrm{S}} : [k_0 + 1] \to \{0, 1\}^*$ for the distribution of $\tilde{K} := \min\{K, k_0 + 1\}$ where $K \sim \mathrm{Zipf}(1 + 1/\ell)$ and k_0 is a large fixed integer (but not too large so it is viable to construct the Shannon code), and then encode $k \in \mathbb{N}^+$ into $f_{\mathrm{S}}(k)$ if $k \leq k_0$, or $f_{\mathrm{S}}(k_0 + 1) \| f_\delta(k - k_0)$ if $k > k_0$, where $f_\delta : \mathbb{N}^+ \to \{0, 1\}^*$ is the Elias delta code, and "$\|$" stands for concatenation.

- A suitable comma code such as the Fibonacci code [88] (which is optimal for a Zipf distribution with a certain parameter).

B

Turning Approximate Markov Chains into Exact Markov Chains

The following lemma shows that if the Markov chain "$X \leftrightarrow U \leftrightarrow Y$" almost holds, that is, there exists random variables $\tilde{X}, \tilde{Y}$ with $\tilde{X} \leftrightarrow U \leftrightarrow \tilde{Y}$ and $\mathbb{P}((X,Y) \neq (\tilde{X},\tilde{Y})) \approx 0$, then there exists a random variable V with small entropy such that $X \leftrightarrow (U,V) \leftrightarrow Y$ holds exactly.

Lemma 76. *For finite discrete random variables $X, Y, \tilde{X}, \tilde{Y}, U$ ($X, \tilde{X} \in \mathcal{X}$ and $Y, \tilde{Y} \in \mathcal{Y}$) with $\tilde{X} \leftrightarrow U \leftrightarrow \tilde{Y}$, there exists a random variable $V \in \mathcal{V}$ with $X \leftrightarrow (U,V) \leftrightarrow Y$, $|\mathcal{V}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}|\} + 1$, and*

$$H(V) \leq H_b(\min\{\eta, 1/2\}) + \eta \log_2 \min\{|\mathcal{X}|, |\mathcal{Y}|\},$$

where H_b is the binary entropy function, and

$$\eta := 2(|\mathcal{X}||\mathcal{Y}|)^{1/4} \sqrt{\mathbb{P}((X,Y) \neq (\tilde{X},\tilde{Y}))}.$$

Proof. We first prove the following claim, which basically states that if two random variables has a small TV distance from being independent, then they are conditionally independent given a random variable that is close to being degenerate:

For finite discrete random variables $X, Y, \tilde{X}, \tilde{Y}$ with $\tilde{X}$ independent of $\tilde{Y}$, there exists a random variable $V \in [0..\min\{|\mathcal{X}|, |\mathcal{Y}|\}]$ with $X \leftrightarrow V \leftrightarrow Y$ and

$$P_V(0) \geq 1 - 2(|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\delta_{\mathrm{TV}}((X,Y),(\tilde{X},\tilde{Y}))}.$$

We now show this claim. Assume $\mathcal{X} = [|\mathcal{X}|]$ and $|\mathcal{X}| \leq |\mathcal{Y}|$. Applying the coupling lemma (Proposition 34), we can assume $\delta_{\mathrm{TV}}((X,Y),(\tilde{X},\tilde{Y})) = \mathbb{P}(E)$, where E is the event $(X,Y) \neq (\tilde{X},\tilde{Y})$. Define $V \in [0..|\mathcal{X}|]$ with

$$\begin{aligned}
&P_{V,X,Y}(0,x,y) \\
&:= \left[P_{\tilde{X}}(x) - \left(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, \tilde{X}=x)}\right]_+ \left[P_{\tilde{Y}}(y) - \left(\frac{|\mathcal{X}|}{|\mathcal{Y}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, \tilde{Y}=y)}\right]_+,
\end{aligned}$$

where $[t]_+ := \max\{t, 0\}$, $P_{V,X,Y}(x,x,y) := P_{X,Y}(x,y) - P_{V,X,Y}(0,x,y)$, and $P_{V,X,Y}(v,x,y) := 0$ for $v \neq x$. To check that this is a valid distribution,

$$\begin{aligned}
&P_{V,X,Y}(0,x,y) \\
&= \left[P_{\tilde{X}}(x) - \left(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, \tilde{X}=x)}\right]_+ \left[P_{\tilde{Y}}(y) - \left(\frac{|\mathcal{X}|}{|\mathcal{Y}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, \tilde{Y}=y)}\right]_+ \\
&\leq \left[P_{\tilde{X}}(x) - \left(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, (\tilde{X},\tilde{Y})=(x,y))}\right]_+ \\
&\qquad \cdot \left[P_{\tilde{Y}}(y) - \left(\frac{|\mathcal{X}|}{|\mathcal{Y}|}\right)^{\frac{1}{4}}\sqrt{\mathbb{P}(E,\, (\tilde{X},\tilde{Y})=(x,y))}\right]_+ \\
&\overset{(a)}{\leq} P_{\tilde{X}}(x)P_{\tilde{Y}}(y) - \mathbb{P}(E,\, (\tilde{X},\tilde{Y})=(x,y)) \\
&= \mathbb{P}((\tilde{X},\tilde{Y})=(x,y)) - \mathbb{P}((X,Y)\neq(x,y),\, (\tilde{X},\tilde{Y})=(x,y)) \\
&\leq P_{X,Y}(x,y),
\end{aligned}$$

where (a) is due to the inequality $[a-s]_+[b-t]_+ \leq [ab-st]_+$ for $a,b,s,t \geq 0$.[1] We have the Markov chain $X \leftrightarrow V \leftrightarrow Y$. We also have

$$\begin{aligned}
\sum_x \left(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\right)^{1/4}\sqrt{\mathbb{P}(E,\, \tilde{X}=x)} &\leq \left(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\right)^{1/4}|\mathcal{X}|\sqrt{\frac{1}{|\mathcal{X}|}\sum_x \mathbb{P}(E,\, \tilde{X}=x)} \\
&= (|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}(E)}.
\end{aligned}$$

[1] Assume $a > s$, $b > t$ (otherwise the inequality is trivial). We have $[a-s]_+[b-t]_+ = ab - at - bs + st < ab - st \leq [ab-st]_+$.

Hence,

$$\begin{aligned}P_V(0) &= \Bigg(\sum_x \bigg[P_{\tilde{X}}(x) - \Big(\frac{|\mathcal{Y}|}{|\mathcal{X}|}\Big)^{1/4}\sqrt{\mathbb{P}(E,\,\tilde{X}=x)}\bigg]_+\Bigg) \\ &\quad\cdot \Bigg(\sum_y \bigg[P_{\tilde{Y}}(y) - \Big(\frac{|\mathcal{X}|}{|\mathcal{Y}|}\Big)^{1/4}\sqrt{\mathbb{P}(E,\,\tilde{Y}=y)}\bigg]_+\Bigg) \\ &\geq \Big[1 - (|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}(E)}\Big]_+^2 \\ &\geq 1 - 2(|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}(E)},\end{aligned}$$

which is the desired claim.

We now prove Lemma 76. Applying the claim on $P_{X,Y,\tilde{X},\tilde{Y}|U}(\cdot|u)$ for each u, there exists $V \in [0..\min\{|\mathcal{X}|,|\mathcal{Y}|\}]$ with $X \leftrightarrow V \leftrightarrow Y$ conditional on $U = u$ (and hence $X \leftrightarrow (U,V) \leftrightarrow Y$) and

$$P_{V|U}(0|u) \geq 1 - 2(|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}((X,Y) \neq (\tilde{X},\tilde{Y})\,|\,U=u)}.$$

We have

$$\begin{aligned}P_V(0) &\geq 1 - \mathbb{E}_U\Big[2(|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}((X,Y)\neq(\tilde{X},\tilde{Y})\,|\,U)}\Big] \\ &\geq 1 - 2(|\mathcal{X}||\mathcal{Y}|)^{1/4}\sqrt{\mathbb{P}((X,Y)\neq(\tilde{X},\tilde{Y}))}.\end{aligned}$$

Hence $P_V(0) \geq 1 - \eta$, and

$$\begin{aligned}H(V) &= H_b(P_V(0)) + (1 - P_V(0))H(V\,|\,V \neq 0) \\ &\leq H_b(\min\{\eta, 1/2\}) + \eta\log_2\min\{|\mathcal{X}|,|\mathcal{Y}|\}.\end{aligned}$$

□

References

[1] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pp. 308–318, 2016.

[2] E. Agustsson and L. Theis, "Universally quantized neural compression," *Advances in neural information processing systems*, vol. 33, 2020, pp. 12 367–12 376.

[3] Y. Altuğ and A. B. Wagner, "Source and channel simulation using arbitrary randomness," *IEEE Transactions on Information Theory*, vol. 58, no. 3, 2012, pp. 1345–1360.

[4] S. Amiri, A. Belloum, S. Klous, and L. Gommans, "Compressive differentially private federated learning through universal vector quantization," in *AAAI Workshop on Privacy-Preserving Artificial Intelligence*, pp. 2–9, 2021.

[5] V. Anantharam and V. Borkar, "Common randomness and distributed control: A counterexample," *Systems & control letters*, vol. 56, no. 7-8, 2007, pp. 568–572.

[6] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pp. 901–914, 2013.

[7] O. Angel and Y. Spinka, "Pairwise optimal coupling of multiple random variables," *arXiv preprint arXiv:1903.00632*, 2019.

[8] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, 2009, pp. 3051–3073.

[9] T. A. Atif, A. Padakandla, and S. S. Pradhan, "Synthesizing correlated randomness using algebraic structured codes," in *2021 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 2417–2422, 2021.

[10] T. A. Atif, A. Padakandla, and S. S. Pradhan, "Source coding for synthesizing correlated randomness," *IEEE Transactions on Information Theory*, vol. 69, no. 1, 2022, pp. 626–649.

[11] J. Ballé, P. A. Chou, D. Minnen, S. Singh, N. Johnston, E. Agustsson, S. J. Hwang, and G. Toderici, "Nonlinear transform coding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 2, 2020, pp. 339–353.

[12] J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," in *5th International Conference on Learning Representations, ICLR 2017*, 2017.

[13] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, and J. A. Hendler, "Towards a theory of semantic communication," in *2011 IEEE Network Science Workshop*, IEEE, pp. 110–117, 2011.

[14] B. Barak, M. Braverman, X. Chen, and A. Rao, "How to compress interactive communication," in *Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 67–76, 2010.

[15] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "On quantum coding for ensembles of mixed states," *Journal of Physics A: Mathematical and General*, vol. 34, no. 35, 2001, p. 6767.

[16] R. Bassily, S. Moran, I. Nachum, J. Shafer, and A. Yehudayoff, "Learners that use little information," in *Algorithmic Learning Theory*, PMLR, pp. 25–55, 2018.

[17] R. Bassily and A. Smith, "Local, private, efficient protocols for succinct histograms," in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pp. 127–135, 2015.

[18] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, 1964, p. 195.

[19] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Transactions on Information Theory*, vol. 60, no. 5, May 2014, pp. 2926–2959. DOI: 10.1109/TIT.2014.2309968.

[20] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Physical review letters*, vol. 70, no. 13, 1993, p. 1895.

[21] C. H. Bennett, P. W. Shor, J. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Transactions on Information Theory*, vol. 48, no. 10, 2002, pp. 2637–2655.

[22] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," *Physical review letters*, vol. 69, no. 20, 1992, p. 2881.

[23] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, NJ, USA, 1971.

[24] T. Berger, "Rate-distortion theory," *Wiley Encyclopedia of Telecommunications*, 2003.

[25] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[26] M. Berta, J. M. Renes, and M. M. Wilde, "Identifying the information gain of a quantum measurement," *IEEE Transactions on Information Theory*, vol. 60, no. 12, 2014, pp. 7987–8006.

[27] Y. Blau and T. Michaeli, "The perception-distortion tradeoff," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6228–6237, 2018.

[28] Y. Blau and T. Michaeli, "Rethinking lossy compression: The rate-distortion-perception tradeoff," in *International Conference on Machine Learning*, PMLR, pp. 675–685, 2019.

[29] M. R. Bloch and J. Kliewer, “Strong coordination over a line network,” in *2013 IEEE International Symposium on Information Theory*, IEEE, pp. 2319–2323, 2013.

[30] A. Block and Y. Polyanskiy, “The sample complexity of approximate rejection sampling with applications to smoothed online learning,” in *The Thirty Sixth Annual Conference on Learning Theory*, PMLR, pp. 228–273, 2023.

[31] S. Bowe, A. Gabizon, and I. Miers, “Scalable multi-party computation for zk-snark parameters in the random beacon model,” *Cryptology ePrint Archive*, 2017.

[32] G. Brassard, R. Cleve, and A. Tapp, “Cost of exactly simulating quantum entanglement with classical communication,” *Physical Review Letters*, vol. 83, no. 9, 1999, p. 1874.

[33] G. Braun, R. Jain, T. Lee, and S. Pokutta, “Information-theoretic approximations of the nonnegative rank,” *computational complexity*, vol. 26, 2017, pp. 147–197.

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

[35] J. Brody, H. Buhrman, M. Kouckỳ, B. Loff, F. Speelman, and N. Vereshchagin, “Towards a reverse Newman’s theorem in interactive information complexity,” *Algorithmica*, vol. 76, 2016, pp. 749–781.

[36] M. Bun, J. Nelson, and U. Stemmer, “Heavy hitters and the structure of local privacy,” *ACM Transactions on Algorithms (TALG)*, vol. 15, no. 4, 2019, pp. 1–40.

[37] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, “Channel simulation: Finite blocklengths and broadcast channels,” *arXiv preprint arXiv:2212.11666*, 2022.

[38] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, “One-shot point-to-point channel simulation,” in *2022 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 796–801, 2022.

[39] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, "Broadcast channel simulation," in *2023 IEEE International Symposium on Information Theory (ISIT)*, pp. 1430–1435, 2023. DOI: 10.1109/ISIT54713.2023.10206649.

[40] N. J. Cerf, N. Gisin, and S. Massar, "Classical teleportation of a quantum bit," *Physical Review Letters*, vol. 84, no. 11, 2000, p. 2521.

[41] G. Cervia, L. Luzzi, M. Le Treust, and M. R. Bloch, "Strong coordination of signals and actions over noisy channels with two-sided state information," *IEEE Transactions on Information Theory*, vol. 66, no. 8, 2020, pp. 4681–4708.

[42] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao, "Informational complexity and the direct sum problem for simultaneous message complexity," in *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, IEEE, pp. 270–278, 2001.

[43] S. Chatterjee and P. Diaconis, "The sample size required in importance sampling," *The Annals of Applied Probability*, vol. 28, no. 2, 2018, pp. 1099–1135.

[44] J. Chen, L. Yu, J. Wang, W. Shi, Y. Ge, and W. Tong, "On the rate-distortion-perception function," *IEEE Journal on Selected Areas in Information Theory*, 2022.

[45] C. F. Choi and C. T. Li, "Multiple-output channel simulation and lossy compression of probability distributions," in *2021 IEEE Information Theory Workshop (ITW)*, IEEE, 2021.

[46] Y. Choi, M. El-Khamy, and J. Lee, "Variable rate deep image compression with a conditional autoencoder," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3146–3154, 2019.

[47] Y. Choi, M. El-Khamy, and J. Lee, "Universal deep neural network compression," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 4, 2020, pp. 715–726.

[48] F. Cicalese, L. Gargano, and U. Vaccaro, "Approximating probability distributions with short vectors, via information theoretic distance measures," in *2016 IEEE ISIT*, IEEE, pp. 1138–1142, 2016.

[49] F. Cicalese, L. Gargano, and U. Vaccaro, "Minimum-entropy couplings and their applications," *IEEE Transactions on Information Theory*, vol. 65, no. 6, 2019, pp. 3436–3451.

[50] F. Cicalese and U. Vaccaro, "Supermodularity and subadditivity properties of the entropy on the majorization lattice," *IEEE Transactions on Information Theory*, vol. 48, no. 4, 2002, pp. 933–938.

[51] J. Clark and U. Hengartner, "On the use of financial data as a random beacon," in *2010 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE 10)*, 2010.

[52] J. E. Cohen and U. G. Rothblum, "Nonnegative ranks, decompositions, and factorizations of nonnegative matrices," *Linear Algebra and its Applications*, vol. 190, 1993, pp. 149–168.

[53] S. Compton, "A tighter approximation guarantee for greedy minimum entropy coupling," in *2022 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 168–173, 2022.

[54] S. Compton, D. Katz, B. Qi, K. Greenewald, and M. Kocaoglu, "Minimum-entropy coupling approximation guarantees beyond the majorization barrier," in *International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 10 445–10 469, 2023.

[55] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, vol. 290. Springer Science & Business Media, 2013.

[56] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[57] T. M. Cover and H. H. Permuter, "Capacity of coordinated actions," in *2007 IEEE International Symposium on Information Theory*, IEEE, pp. 2701–2705, 2007.

[58] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.

[59] J. A. Csirik, "Cost of exactly simulating a bell pair using classical communication," *Physical Review A*, vol. 66, no. 1, 2002, p. 014 302.

[60] I. Csiszár, "The method of types," *IEEE Transactions on Information Theory*, vol. 44, no. 6, 1998, pp. 2505–2523.

[61] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[62] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, "Zero-error channel capacity and simulation assisted by non-local correlations," *IEEE Transactions on Information Theory*, vol. 57, no. 8, 2011, pp. 5509–5523.

[63] P. Cuff, "Communication requirements for generating correlated random variables," in *2008 IEEE International Symposium on Information Theory*, IEEE, pp. 1393–1397, 2008.

[64] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, Nov. 2013, pp. 7071–7096.

[65] P. Cuff, "Soft covering with high probability," in *2016 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 2963–2967, 2016.

[66] P. Cuff, H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 9, Sep. 2010, pp. 4181–4206.

[67] L. Devroye, *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.

[68] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *2013 IEEE 54th annual symposium on foundations of computer science*, IEEE, pp. 429–438, 2013.

[69] E. Dupont, H. Loya, M. Alizadeh, A. Golinski, Y. Teh, and A. Doucet, "COIN++: Neural compression across modalities," *Transactions on Machine Learning Research*, vol. 2022, no. 11, 2022.

[70] E. Dupont, A. Golinski, M. Alizadeh, Y. W. Teh, and A. Doucet, "COIN: COmpression with Implicit Neural representations," in *Neural Compression: From Information Theory to Applications–Workshop@ ICLR 2021*, 2021.

[71] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of cryptography conference*, Springer, pp. 265–284, 2006.

[72] C. Dwork, A. Roth, *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, 2014, pp. 211–407.

[73] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[74] P. Elias, "The efficient construction of an unbiased random sequence," *The Annals of Mathematical Statistics*, vol. 43, no. 3, 1972, pp. 865–870.

[75] P. Elias, "Universal codeword sets and representations of the integers," *IEEE Transactions on Information Theory*, vol. 21, no. 2, 1975, pp. 194–203.

[76] E. Erdemir, T.-Y. Tung, P. L. Dragotti, and D. Gündüz, "Generative joint source-channel coding for semantic image transmission," *IEEE Journal on Selected Areas in Communications*, 2023.

[77] A. Evfimievski, J. Gehrke, and R. Srikant, "Limiting privacy breaches in privacy preserving data mining," in *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 211–222, 2003.

[78] R. M. Fano, *The transmission of information*, vol. 65. Massachusetts Institute of Technology, Research Laboratory of Electronics, 1949.

[79] R. M. Fano, *Transmission of Information*. The MIT Press, 1961.

[80] V. Feldman and K. Talwar, "Lossless compression of efficient private local randomizers," in *International Conference on Machine Learning*, PMLR, pp. 3208–3219, 2021.

[81] M. Feldmann, "New loophole for the einstein-podolsky-rosen paradox," *Foundations of Physics Letters*, vol. 8, 1995, pp. 41–53.

[82] G. Flamich, "Greedy Poisson rejection sampling," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[83] G. Flamich, M. Havasi, and J. M. Hernández-Lobato, "Compressing images by encoding their latent representations with relative entropy coding," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 16 131–16 141.

[84] G. Flamich, S. Markou, and J. M. Hernández-Lobato, "Fast relative entropy coding with A* coding," in *International Conference on Machine Learning*, PMLR, pp. 6548–6577, 2022.

[85] G. Flamich, S. Markou, and J. M. Hernández-Lobato, "Faster relative entropy coding with greedy rejection coding," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[86] G. Flamich and L. Theis, "Adaptive greedy rejection sampling," in *2023 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 454–459, 2023.

[87] G. Flamich and L. Wells, "Some notes on the sample complexity of approximate channel simulation," *arXiv preprint arXiv:2405.04363*, 2024.

[88] A. S. Fraenkel and S. T. Kleinb, "Robust universal complete codes for transmission and compression," *Discrete Applied Mathematics*, vol. 64, no. 1, 1996, pp. 31–55.

[89] P. E. Frenkel and M. Weiner, "Classical information storage in an n-level quantum system," *Communications in Mathematical Physics*, vol. 340, no. 2, 2015, pp. 563–574.

[90] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, no. 5, 1968, pp. 676–683.

[91] D. Goc and G. Flamich, "On channel simulation with causal rejection samplers," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 1682–1687, 2024. DOI: 10.1109/ISIT57864.2024.10619339.

[92] A. Gohari, M. H. Yassaee, and M. R. Aref, "Secure channel simulation," in *2012 IEEE Information Theory Workshop*, IEEE, pp. 406–410, 2012.

[93] A. A. Gohari and V. Anantharam, "Generating dependent random variables over networks," in *2011 IEEE Information Theory Workshop*, IEEE, pp. 698–702, 2011.

[94] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, 2006, pp. 1603–1636.

[95] R. M. Gray and T. G. Stockham, “Dithered quantizers,” *IEEE Transactions on Information Theory*, vol. 39, no. 3, 1993, pp. 805–812.

[96] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, 1998, pp. 2325–2383.

[97] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, “Beyond transmitting bits: Context, semantics, and task-oriented communications,” *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, 2022, pp. 5–41.

[98] Z. Guo, G. Flamich, J. He, Z. Chen, and J. M. Hernández-Lobato, “Compression with Bayesian implicit neural representations,” *Advances in Neural Information Processing Systems*, vol. 36, 2023, pp. 1938–1956.

[99] F. Haddadpour, M. H. Yassaee, S. Beigi, A. Gohari, and M. R. Aref, “Simulation of a channel with another channel,” *IEEE Transactions on Information Theory*, vol. 63, no. 5, 2016, pp. 2659–2677.

[100] B. Hajek and M. Pursley, “Evaluation of an achievable rate region for the broadcast channel,” *IEEE Transactions on Information Theory*, vol. 25, no. 1, Jan. 1979, pp. 36–46. DOI: 10.1109/TIT.1979.1055989.

[101] Y. Hamdi, A. B. Wagner, and D. Gündüz, “The rate-distortion-perception trade-off: The role of private randomness,” *arXiv preprint arXiv:2404.01111*, 2024.

[102] T. S. Han and S. Verdú, “Approximation theory of output statistics,” *IEEE Transactions on Information Theory*, vol. 39, no. 3, May 1993, pp. 752–772. DOI: 10.1109/18.256486.

[103] T. S. Han and M. Hoshi, “Interval algorithm for random number generation,” *IEEE Transactions on Information Theory*, vol. 43, no. 2, Mar. 1997, pp. 599–611. DOI: 10.1109/18.556116.

[104] H. Haramoto, M. Matsumoto, and P. L’Ecuyer, “A fast jump ahead algorithm for linear recurrences in a polynomial space,” in *International Conference on Sequences and Their Applications*, Springer, pp. 290–298, 2008.

[105] H. Haramoto, M. Matsumoto, T. Nishimura, F. Panneton, and P. L'Ecuyer, "Efficient jump ahead for F2-linear random number generators," *INFORMS Journal on Computing*, vol. 20, no. 3, 2008, pp. 385–390.

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

[107] B. Hasırcıoğlu and D. Gündüz, "Communication efficient private federated learning using dithering," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 7575–7579, 2024.

[108] M. Havasi, R. Peharz, and J. M. Hernández-Lobato, "Minimal random code learning: Getting bits back from compressed model parameters," in *7th International Conference on Learning Representations, ICLR 2019*, 2019.

[109] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 52, no. 4, 2006, pp. 1562–1575.

[110] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Transactions on Information Theory*, vol. 54, no. 10, 2008, pp. 4619–4637.

[111] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, 2009, pp. 4947–4966.

[112] J. He, G. Flamich, Z. Guo, and J. M. Hernández-Lobato, "Recombiner: Robust and enhanced compression with bayesian implicit neural representations," in *The Twelfth International Conference on Learning Representations*, 2024.

[113] J. He, G. Flamich, and J. M. Hernández-Lobato, "Accelerating relative entropy coding with space partitioning," *arXiv preprint arXiv:2405.12203*, 2024.

[114] M. Hegazy, R. Leluc, C. T. Li, and A. Dieuleveut, “Compression with exact error distribution for federated learning,” in *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, vol. 238, pp. 613–621, PMLR, Feb. 2024.

[115] M. Hegazy and C. T. Li, “Randomized quantization with exact error distribution,” in *2022 IEEE Information Theory Workshop (ITW)*, IEEE, pp. 350–355, 2022.

[116] W. Hoeffding and G. Simons, “Unbiased coin tossing with a biased coin,” in *The Collected Works of Wassily Hoeffding*, Springer, 1970, pp. 501–512.

[117] A. S. Holevo, “Bounds for the quantity of information transmitted by a quantum communication channel,” *Problemy Peredachi Informatsii*, vol. 9, no. 3, 1973, pp. 3–11.

[118] D. A. Huffman, “A method for the construction of minimum-redundancy codes,” *Proceedings of the IRE*, vol. 40, no. 9, 1952, pp. 1098–1101.

[119] I. A. Huijben, W. Kool, M. B. Paulus, and R. J. Van Sloun, “A review of the gumbel-max trick and its extensions for discrete stochasticity in machine learning,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 2, 2022, pp. 1353–1371.

[120] B. Isik, F. Pase, D. Gunduz, S. Koyejo, T. Weissman, and M. Zorzi, “Adaptive compression in federated learning via side information,” in *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, S. Dasgupta, S. Mandt, and Y. Li, Eds., ser. Proceedings of Machine Learning Research, vol. 238, pp. 487–495, PMLR, Feb. 2024.

[121] K. Itô, *An Introduction to Probability Theory*. Cambridge University Press, 1984.

[122] R. Jain, J. Radhakrishnan, and P. Sen, “A direct sum theorem in communication complexity via message compression,” in *Automata, Languages and Programming: 30th International Colloquium, ICALP 2003 Eindhoven, The Netherlands, June 30–July 4, 2003 Proceedings 30*, Springer, pp. 300–315, 2003.

[123] R. Jain, Y. Shi, Z. Wei, and S. Zhang, "Efficient protocols for generating bipartite classical distributions and quantum states," *IEEE Transactions on Information Theory*, vol. 59, no. 8, 2013, pp. 5171–5178.

[124] N. Jayant and L. Rabiner, "The application of dither to the quantization of speech signals," *Bell System Technical Journal*, vol. 51, no. 6, 1972, pp. 1293–1304.

[125] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE transactions on communication technology*, vol. 15, no. 1, 1967, pp. 52–60.

[126] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, *et al.*, "Advances and open problems in federated learning," *Foundations and trends® in machine learning*, vol. 14, no. 1–2, 2021, pp. 1–210.

[127] O. Kallenberg, *Foundations of Modern Probability*. Springer Science & Business Media, 2002.

[128] S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith, "What can we learn privately?" *SIAM Journal on Computing*, vol. 40, no. 3, 2011, pp. 793–826.

[129] J. Kemperman, "On the shannon capacity of an arbitrary channel," in *Indagationes Mathematicae (Proceedings)*, North-Holland, vol. 77, pp. 101–115, 1974.

[130] A. Khisti, A. Behboodi, G. Cesa, and P. Kumar, "Unequal message protection: One-shot analysis via Poisson matching lemma," in *2024 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 629–634, 2024.

[131] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[132] A. Kirac and P. Vaidyanathan, "Results on lattice vector quantization with dithering," *IEEE Transactions On Circuits and Systems II: Analog and Digital Signal Processing*, vol. 43, no. 12, 1996, pp. 811–826.

[133] T. Kloek and H. K. Van Dijk, "Bayesian estimates of equation system parameters: An application of integration by monte carlo," *Econometrica: Journal of the Econometric Society*, 1978, pp. 1–19.

[134] R. T. Kneusel, *Random numbers and computers*, vol. 239. Springer, 2018.

[135] D. E. Knuth and A. C. Yao, "The complexity of nonuniform random number generation," *Algorithms and Complexity: New Directions and Recent Results*, J. F. Traub, Ed., 1976, pp. 357–428.

[136] S. Kobus, L. Theis, and D. Gündüz, "Gaussian channel simulation with rotated dithered quantization," in *2024 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 1907–1912, 2024.

[137] S. Kobus, T.-Y. Tung, and D. Gündüz, "Universal sample coding," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[138] M. Kocaoglu, A. G. Dimakis, S. Vishwanath, and B. Hassibi, "Entropic causal inference," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[139] M. Kocaoglu, A. G. Dimakis, S. Vishwanath, and B. Hassibi, "Entropic causality and greedy minimum entropy coupling," in *2017 IEEE ISIT*, IEEE, pp. 1465–1469, 2017.

[140] A. N. Kolmogorov, "On certain asymptotic characteristics of completely bounded metric spaces," in *Dokl. Akad. Nauk SSSR*, vol. 108, pp. 385–388, 1956.

[141] M. Kovačević, I. Stanojević, and V. Šenk, "On the entropy of couplings," *Information and Computation*, vol. 242, 2015, pp. 369–382.

[142] L. G. Kraft, "A device for quantizing, grouping, and coding amplitude-modulated pulses," M.S. thesis, Massachusetts Institute of Technology, 1949.

[143] G. Kramer and S. A. Savari, "Communicating probability distributions," *IEEE Transactions on Information Theory*, vol. 53, no. 2, 2007, pp. 518–525.

[144] G. R. Kumar, C. T. Li, and A. El Gamal, "Exact common information," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 161–165, Jun. 2014.

[145] G. R. Kurri, V. M. Prabhakaran, and A. D. Sarwate, "Coordination through shared randomness," *IEEE Transactions on Information Theory*, vol. 67, no. 8, 2021, pp. 4948–4974.

[146] G. R. Kurri, V. Ramachandran, S. R. B. Pillai, and V. M. Prabhakaran, "Multiple access channel simulation," *IEEE Transactions on Information Theory*, vol. 68, no. 11, 2022, pp. 7575–7603.

[147] N. Lang, E. Sofer, T. Shaked, and N. Shlezinger, "Joint privacy enhancement and quantization in federated learning," *IEEE Transactions on Signal Processing*, vol. 71, 2023, pp. 295–310.

[148] G. Last and M. Penrose, *Lectures on the Poisson Process*, vol. 7. Cambridge University Press, 2017.

[149] M. Le Treust and T. Tomala, "Strategic coordination with state information at the decoder," in *International Zurich Seminar on Information and Communication (IZS 2018). Proceedings*, ETH Zurich, pp. 30–34, 2018.

[150] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, 1999, pp. 788–791.

[151] S.-H. Lee and S.-Y. Chung, "A unified approach for network information theory," in *2015 IEEE ISIT*, IEEE, pp. 1277–1281, 2015.

[152] S.-H. Lee and S.-Y. Chung, "A unified random coding bound," *IEEE Transactions on Information Theory*, vol. 64, no. 10, 2018, pp. 6779–6802.

[153] E. Lei, H. Hassani, and S. S. Bidokhti, "Neural estimation of the rate-distortion function with applications to operational source coding," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 4, 2022, pp. 674–686.

[154] C. T. Li and A. El Gamal, "A universal coding scheme for remote generation of continuous random variables," *IEEE Transactions on Information Theory*, vol. 64, no. 4, Apr. 2018, pp. 2583–2592. DOI: 10.1109/TIT.2018.2803752.

[155] C. T. Li and A. El Gamal, "Strong functional representation lemma and applications to coding theorems," *IEEE Transactions on Information Theory*, vol. 64, no. 11, Nov. 2018, pp. 6967–6978. DOI: 10.1109/TIT.2018.2865570.

[156] C. T. Li, X. Wu, A. Ozgur, and A. El Gamal, "Minimax learning for remote prediction," in *2018 IEEE ISIT*, pp. 541–545, Jun. 2018. DOI: 10.1109/ISIT.2018.8437318.

[157] C. T. Li, *PSITIP - Python symbolic information theoretic inequality prover*, 2020. URL: https://github.com/cheuktingli/psitip.

[158] C. T. Li, "Efficient approximate minimum entropy coupling of multiple probability distributions," *IEEE Transactions on Information Theory*, vol. 67, no. 8, 2021, pp. 5259–5268. DOI: 10.1109/TIT.2021.3076986.

[159] C. T. Li, "An automated theorem proving framework for information-theoretic results," *IEEE Transactions on Information Theory*, vol. 69, no. 11, 2023, pp. 6857–6877. DOI: 10.1109/TIT.2023.3296597.

[160] C. T. Li, "Pointwise redundancy in one-shot lossy compression via Poisson functional representation," in *arXiv preprint; short version presented at 2024 International Zurich Seminar on Information and Communication*, pp. 28–29, 2024. URL: https://arxiv.org/pdf/2401.14805.pdf.

[161] C. T. Li and V. Anantharam, "Pairwise multi-marginal optimal transport and embedding for earth mover's distance," *arXiv preprint arXiv:1908.01388*, 2019.

[162] 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, 2021, pp. 2624–2651.

[163] C. T. Li and A. El Gamal, "Distributed simulation of continuous random variables," *IEEE Transactions on Information Theory*, vol. 63, no. 10, 2017, pp. 6329–6343.

[164] M. Li, J. Klejsa, and W. B. Kleijn, "Distribution preserving quantization with dithering and transformation," *IEEE Signal Processing Letters*, vol. 17, no. 12, 2010, pp. 1014–1017.

[165] M. Li, J. Klejsa, and W. B. Kleijn, "On distribution preserving quantization," *arXiv preprint arXiv:1108.3728*, 2011.

[166] T. Lindvall, *Lectures on the coupling method.* Courier Corporation, 2002.

[167] C. W. Ling and C. T. Li, "Vector quantization with error uniformly distributed over an arbitrary set," in *2023 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 856–861, 2023.

[168] C. W. Ling and C. T. Li, "Rejection-sampled universal quantization for smaller quantization errors," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 1883–1888, 2024. DOI: 10.1109/ISIT57864.2024.10619181.

[169] J. Liu and S. Verdú, "Rejection sampling and noncausal sampling under moment constraints," in *2018 IEEE ISIT*, pp. 1565–1569, Jun. 2018. DOI: 10.1109/ISIT.2018.8437857.

[170] J. Liu, M. H. Yassaee, and S. Verdú, "Sharp bounds for mutual covering," *IEEE Transactions on Information Theory*, vol. 65, no. 12, 2019, pp. 8067–8083.

[171] J. S. Liu, *Monte Carlo Strategies in Scientific Computing.* New York, NY, USA: Springer, 2004.

[172] W. Liu, G. Xu, and B. Chen, "The common information of N dependent random variables," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, pp. 836–843, 2010.

[173] Y. Liu, W.-N. Chen, A. Özgür, and C. T. Li, "Universal exact compression of differentially private mechanisms," *Thirty-Eighth Annual Conference on Neural Information Processing Systems (NeurIPS 2024)*, 2024.

[174] Y. Liu and C. T. Li, "One-shot coding over general noisy networks," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 3124–3129, 2024.

[175] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Transactions on Information Theory*, vol. 57, no. 9, 2011, pp. 6180–6195.

[176] D. J. MacKay, *Information theory, inference and learning algorithms.* Cambridge University Press, 2003.

[177] C. J. Maddison, "A Poisson process model for Monte Carlo," *Perturbation, Optimization, and Statistics*, 2016, pp. 193–232.
[178] C. J. Maddison, D. Tarlow, and T. Minka, "A* sampling," *Advances in neural information processing systems*, vol. 27, 2014.
[179] M. A. Managoli and V. M. Prabhakaran, "Broadcast channel synthesis from shared randomness," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 1919–1924, 2024. DOI: 10.1109/ISIT57864.2024.10619094.
[180] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of Majorization and its Applications*. New York, Dordrecht, Heidelberg, London: Springer, 2011.
[181] S. Massar, D. Bacon, N. J. Cerf, and R. Cleve, "Classical simulation of quantum entanglement without local hidden variables," *Physical Review A*, vol. 63, no. 5, 2001, p. 052 305.
[182] R. Matsumoto, "Introducing the perception-distortion tradeoff into the rate-distortion theory of general information sources," *IEICE Communications Express*, vol. 7, no. 11, 2018, pp. 427–431.
[183] R. Matsumoto, "Rate-distortion-perception tradeoff of variable-length source coding for general information sources," *IEICE Communications Express*, vol. 8, no. 2, 2019, pp. 38–42.
[184] T. Maudlin, "Bell's inequality, information transmission, and prism models," in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Cambridge University Press, vol. 1992, pp. 404–417, 1992.
[185] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, PMLR, pp. 1273–1282, 2017.
[186] H. B. McMahan, D. Ramage, K. Talwar, and L. Zhang, "Learning differentially private recurrent language models," in *International Conference on Learning Representations*, 2018.

[187] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, and S. Rangan, "5G mmWave module for the ns-3 network simulator," in *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 283–290, 2015.

[188] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, 2021, pp. 99–106.

[189] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, 2017.

[190] A. Nema, S. Sreekumar, and M. Berta, "One-shot multiple access channel simulation," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 2981–2986, 2024. DOI: 10.1109/ISIT57864.2024.10619283.

[191] NumPy Developers, *Mersenne twister (mt19937)*, 2024. URL: https://numpy.org/doc/stable/reference/random/bit_generators/mt19937.html.

[192] M. E. O'Neill, "PCG: A family of simple fast space-efficient statistically good algorithms for random number generation," *ACM Transactions on Mathematical Software*, 2014.

[193] S. A. Obead, B. N. Vellambi, and J. Kliewer, "Strong coordination over noisy channels," *IEEE Transactions on Information Theory*, vol. 67, no. 5, 2021, pp. 2716–2738.

[194] G. Oded, *Foundations of Cryptography: Basic Tools*. Cambridge: Cambridge university press, 2004.

[195] Y. Oohama, "Performance analysis of the interval algorithm for random number generation based on number systems," *IEEE Transactions on Information Theory*, vol. 57, no. 3, 2011, pp. 1177–1185.

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

[197] A. Painsky, S. Rosset, and M. Feder, "Memoryless representation of Markov processes," in *2013 IEEE ISIT*, IEEE, pp. 2294–298, 2013.

[198] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "DeepSDF: Learning continuous signed distance functions for shape representation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 165–174, 2019.

[199] F. Pase, S. Kobus, D. Gündüz, and M. Zorzi, "Semantic communication of learnable concepts," in *2023 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 731–736, 2023.

[200] Y. Peres, "Iterating von Neumann's procedure for extracting random bits," *The Annals of Statistics*, 1992, pp. 590–597.

[201] B. Phan, A. Khisti, and C. Louizos, "Importance matching lemma for lossy compression with side information," in *International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 1387–1395, 2024.

[202] S. Pirandola, S. L. Braunstein, R. Laurenza, C. Ottaviani, T. P. Cope, G. Spedalieri, and L. Banchi, "Theory of channel simulation and bounds for private communication," *Quantum Science and Technology*, vol. 3, no. 3, 2018, p. 035 009.

[203] 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, 2010, pp. 2307–2359.

[204] Y. Polyanskiy and Y. Wu, *Information theory: From coding to learning*. Cambridge University Press, 2024.

[205] J. Propp and D. Wilson, "Coupling from the past: A user's guide," *Microsurveys in Discrete Probability*, vol. 41, 1998, pp. 181–192.

[206] J. G. Propp and D. B. Wilson, "Exact sampling with coupled Markov chains and applications to statistical mechanics," *Random Structures & Algorithms*, vol. 9, no. 1-2, 1996, pp. 223–252.

[207] M. O. Rabin, "Transaction protection by beacons," *Journal of Computer and System Sciences*, vol. 27, no. 2, 1983, pp. 256–267.

[208] V. Ramachandran, T. J. Oechtering, and M. Skoglund, "Multiterminal strong coordination over noiseless networks with secrecy constraints," in *International Zurich Seminar on Information and Communication (IZS 2024). Proceedings*, ETH Zürich, pp. 159–163, 2024.

[209] RAND Corporation, *A Million Random Digits with 100,000 Normal Deviates*. Santa Monica, CA: RAND Corporation, 2001. DOI: 10.7249/MR1418.

[210] A. Rao and A. Yehudayoff, *Communication Complexity and Applications*. Cambridge University Press, 2020.

[211] R. Renner and S. Wolf, "New bounds in secret-key agreement: The gap between formation and secrecy extraction," in *Advances in Cryptology–EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland*, Springer, pp. 562–577, 2003.

[212] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, The Regents of the University of California, 1961.

[213] L. Roberts, "Picture coding using pseudo-random noise," *IRE Transactions on Information Theory*, vol. 8, no. 2, 1962, pp. 145–154.

[214] J. R. Roche, "Efficient generation of random variables from biased coins," in *Proceedings. 1991 IEEE International Symposium on Information Theory*, IEEE, pp. 169–169, 1991.

[215] S. Ross, *A First Course in Probability*. Pearson Higher Ed, 2019.

[216] M. Rossi, "Greedy additive approximation algorithms for minimum-entropy coupling problem," in *2019 IEEE ISIT*, IEEE, pp. 1127–1131, 2019.

[217] N. Saldi, T. Linder, and S. Yüksel, "Randomized quantization and optimal design with a marginal constraint," in *2013 IEEE International Symposium on Information Theory*, IEEE, pp. 2349–2353, 2013.

[218] N. Saldi, T. Linder, and S. Yüksel, "Randomized quantization and source coding with constrained output distribution," *IEEE Transactions on Information Theory*, vol. 61, no. 1, 2014, pp. 91–106.

[219] N. Saldi, T. Linder, and S. Yüksel, "Output constrained lossy source coding with limited common randomness," *IEEE Transactions on Information Theory*, vol. 61, no. 9, 2015, pp. 4984–4998.

[220] J. K. Salmon, M. A. Moraes, R. O. Dror, and D. E. Shaw, "Parallel random numbers: As easy as 1, 2, 3," in *Proceedings of 2011 international conference for high performance computing, networking, storage and analysis*, pp. 1–12, 2011.

[221] S. Satpathy and P. Cuff, "Secure coordination with a two-sided helper," in *2014 IEEE International Symposium on Information Theory*, IEEE, pp. 406–410, 2014.

[222] S. Satpathy and P. Cuff, "Secure cascade channel synthesis," *IEEE Transactions on Information Theory*, vol. 62, no. 11, 2016, pp. 6081–6094.

[223] K. Sayood, *Introduction to data compression*. Morgan Kaufmann, 2018.

[224] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Transactions on Communication Technology*, vol. 12, no. 4, 1964, pp. 162–165.

[225] M. Sefidgaran, A. Zaidi, and P. Krasnowski, "Minimal communication-cost statistical learning," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 777–782, 2024. DOI: 10.1109/ISIT57864.2024.10619653.

[226] A. Shah, W.-N. Chen, J. Balle, P. Kairouz, and L. Theis, "Optimal compression of locally differentially private mechanisms," in *International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 7680–7723, 2022.

[227] A. M. Shahmiri, C. W. Ling, and C. T. Li, "Communication-efficient Laplace mechanism for differential privacy via random quantization," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 4550–4554, 2024.

[228] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, 1948, pp. 379–423.
[229] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, 1949, pp. 656–715.
[230] Y. Shao, Q. Cao, and D. Gunduz, "A theory of semantic communication," *arXiv preprint arXiv:2212.01485*, 2022.
[231] Y. Shkel, "Functional representation lemma: Algorithms and applications," in *International Zurich Seminar on Information and Communication (IZS 2024)*, p. 26, 2024.
[232] Y. Y. Shkel and A. K. Yadav, "Information spectrum converse for minimum entropy couplings and functional representations," in *2023 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 66–71, 2023.
[233] N. Shlezinger, M. Chen, Y. C. Eldar, H. V. Poor, and S. Cui, "UVeQFed: Universal vector quantization for federated learning," *IEEE Transactions on Signal Processing*, vol. 69, 2020, pp. 500–514.
[234] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 7462–7473.
[235] E. C. Song, P. Cuff, and H. V. Poor, "The likelihood encoder for lossy compression," *IEEE Transactions on Information Theory*, vol. 62, no. 4, 2016, pp. 1836–1849.
[236] S. M. Sriramu, R. Barsz, E. Polito, and A. B. Wagner, "Fast channel simulation via error-correcting codes," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.
[237] S. M. Sriramu and A. B. Wagner, "Optimal redundancy in exact channel synthesis," in *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 1913–1918, 2024. DOI: 10.1109/ISIT57864.2024.10619703.
[238] K. O. Stanley, "Compositional pattern producing networks: A novel abstraction of development," *Genetic programming and evolvable machines*, vol. 8, 2007, pp. 131–162.

[239] Y. Steinberg and S. Verdú, "Channel simulation and coding with side information," *IEEE Transactions on Information Theory*, vol. 40, no. 3, 1994, pp. 634–646.
[240] Y. Steinberg and S. Verdú, "Simulation of random processes and rate-distortion theory," *IEEE Transactions on Information Theory*, vol. 42, no. 1, 1996, pp. 63–86.
[241] M. Steiner, "Towards quantifying non-local information transfer: Finite-bit non-locality," *Physics Letters A*, vol. 270, no. 5, 2000, pp. 239–244.
[242] S. Sun, G. R. MacCartney, and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5g wireless communications," in *2017 IEEE international conference on communications (ICC)*, IEEE, pp. 1–7, 2017.
[243] M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 7537–7547.
[244] L. Theis, T. Salimans, M. D. Hoffman, and F. Mentzer, "Lossy compression with Gaussian diffusion," *arXiv preprint arXiv:2206.08889*, 2022.
[245] L. Theis and A. B. Wagner, "A coding theorem for the rate-distortion-perception function," in *Neural Compression: From Information Theory to Applications–Workshop@ ICLR 2021*, 2021.
[246] L. Theis and N. Yosri, "Algorithms for the communication of samples," in *International Conference on Machine Learning*, PMLR, pp. 21 308–21 328, 2022.
[247] A. Triastcyn, M. Reisser, and C. Louizos, "DP-REC: Private & communication-efficient federated learning," *arXiv preprint arXiv:2111.05454*, 2021.
[248] M. Tschannen, E. Agustsson, and M. Lucic, "Deep generative models for distribution-preserving lossy compression," *Advances in neural information processing systems*, vol. 31, 2018.

[249] T. Uyematsu and F. Kanaya, "Channel simulation by interval algorithm: A performance analysis of interval algorithm," *IEEE Transactions on Information Theory*, vol. 45, no. 6, 1999, pp. 2121–2129.

[250] A. Vandaele, N. Gillis, F. Glineur, and D. Tuyttens, "Heuristics for exact nonnegative matrix factorization," *Journal of Global Optimization*, vol. 65, 2016, pp. 369–400.

[251] S. A. Vavasis, "On the complexity of nonnegative matrix factorization," *SIAM Journal on Optimization*, vol. 20, no. 3, 2010, pp. 1364–1377.

[252] B. N. Vellambi, J. Kliewer, and M. R. Bloch, "Strong coordination over multi-hop line networks using channel resolvability codebooks," *IEEE Transactions on Information Theory*, vol. 64, no. 2, 2017, pp. 1132–1162.

[253] B. N. Vellambi and J. Kliewer, "New results on the equality of exact and wyner common information rates," in *2018 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 151–155, 2018.

[254] M. Vidyasagar, "A metric between probability distributions on finite sets of different cardinalities and applications to order reduction," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, 2012, pp. 2464–2477.

[255] J. Von Neumann, "Various techniques used in connection with random digits," *John von Neumann, Collected Works*, vol. 5, 1963, pp. 768–770.

[256] A. B. Wagner, "The rate-distortion-perception tradeoff: The role of common randomness," *arXiv preprint arXiv:2202.04147*, 2022.

[257] S. Walker, "The uniform power distribution," *Journal of Applied Statistics*, vol. 26, no. 4, 1999, pp. 509–517.

[258] R. A. Wannamaker, S. P. Lipshitz, J. Vanderkooy, and J. N. Wright, "A theory of nonsubtractive dither," *IEEE Transactions on Signal Processing*, vol. 48, no. 2, 2000, pp. 499–516.

[259] S. Watanabe, S. Kuzuoka, and V. Y. F. Tan, "Nonasymptotic and second-order achievability bounds for coding with side-information," *IEEE Transactions on Information Theory*, vol. 61, no. 4, Apr. 2015, pp. 1574–1605. DOI: 10.1109/TIT.2015.2400994.

[260] S. Watanabe and T. S. Han, "Interval algorithm for random number generation: Information spectrum approach," *IEEE Transactions on Information Theory*, vol. 66, no. 3, 2020, pp. 1691–1701. DOI: 10.1109/TIT.2019.2946235.

[261] S. Watanabe and M. Hayashi, "Strong converse and second-order asymptotics of channel resolvability," in *2014 IEEE International Symposium on Information Theory*, IEEE, pp. 1882–1886, 2014.

[262] W. Weaver, "Recent contributions to the mathematical theory of communication," *ETC: a Review of General Semantics*, 1953, pp. 261–281.

[263] T. Weissman and E. Ordentlich, "The empirical distribution of rate-constrained source codes," *IEEE Transactions on Information Theory*, vol. 51, no. 11, 2005, pp. 3718–3733.

[264] F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Transactions on Information Theory*, vol. 31, no. 3, May 1985, pp. 313–327. DOI: 10.1109/TIT.1985.1057042.

[265] D. B. Wilson, "Layered multishift coupling for use in perfect sampling algorithms (with a primer on CFTP)," *Monte Carlo Methods*, vol. 26, 2000, pp. 141–176.

[266] A. Winter, "Compression of sources of probability distributions and density operators," *arXiv preprint quant-ph/0208131*, 2002.

[267] A. Wyner and J. Ziv, "The sliding-window Lempel-Ziv algorithm is asymptotically optimal," *Proceedings of the IEEE*, vol. 82, no. 6, 1994, pp. 872–877. DOI: 10.1109/5.286191.

[268] A. D. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, 1975, pp. 163–179.

[269] A. D. Wyner, "The wire-tap channel," *Bell system technical journal*, vol. 54, no. 8, 1975, pp. 1355–1387.

[270] S. Yagli and P. Cuff, "Exact exponent for soft covering," *IEEE Transactions on Information Theory*, vol. 65, no. 10, 2019, pp. 6234–6262.

[271] G. Yan, T. Li, T. Lan, K. Wu, and L. Song, "Layered randomized quantization for communication-efficient and privacy-preserving distributed learning," *arXiv preprint arXiv:2312.07060*, 2023.

[272] Y. Yang, R. Bamler, and S. Mandt, "Improving inference for neural image compression," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 573–584.

[273] Y. Yang, S. Mandt, and L. Theis, "An introduction to neural data compression," *Foundations and Trends® in Computer Graphics and Vision*, vol. 15, no. 2, 2023, pp. 113–200.

[274] A. C.-C. Yao, "Some complexity questions related to distributive computing," in *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pp. 209–213, 1979.

[275] M. H. Yassaee, M. R. Aref, and A. Gohari, "Non-asymptotic output statistics of random binning and its applications," in *2013 IEEE ISIT*, pp. 1849–1853, Jul. 2013. DOI: 10.1109/ISIT.2013.6620547.

[276] M. H. Yassaee, "Almost exact analysis of soft covering lemma via large deviation," in *2019 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 1387–1391, 2019.

[277] M. H. Yassaee, "One-shot achievability via fidelity," in *Proc. IEEE Int. Symp. Inf. Theory*, IEEE, pp. 301–305, 2015.

[278] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability proof via output statistics of random binning," *IEEE Transactions on Information Theory*, vol. 60, no. 11, 2014, pp. 6760–6786.

[279] M. H. Yassaee, A. Gohari, and M. R. Aref, "Channel simulation via interactive communications," *IEEE Transactions on Information Theory*, vol. 61, no. 6, 2015, pp. 2964–2982.

[280] R. W. Yeung, *Information theory and network coding*. New York: Springer Science & Business Media, 2008.

[281] L. Yu and V. Y. Tan, "Asymptotic coupling and its applications in information theory," *IEEE Transactions on Information Theory*, vol. 65, no. 3, 2018, pp. 1321–1344.

[282] L. Yu and V. Y. Tan, "Exact channel synthesis," *IEEE Transactions on Information Theory*, vol. 66, no. 5, 2019, pp. 2799–2818.

[283] L. Yu and V. Y. Tan, "On exact and ∞-rényi common informations," *IEEE Transactions on Information Theory*, vol. 66, no. 6, 2020, pp. 3366–3406.

[284] L. Yu and V. Y. Tan, "Common information, noise stability, and their extensions," *Foundations and Trends® in Communications and Information Theory*, vol. 19, no. 2, 2022, pp. 107–389.

[285] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory*. Cambridge University Press, 2014. DOI: 10.1017/CBO9781139045520.

[286] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, vol. 38, no. 2, 1992, pp. 428–436.

[287] G. Zhang, J. Qian, J. Chen, and A. Khisti, "Universal rate-distortion-perception representations for lossy compression," *Advances in Neural Information Processing Systems*, vol. 34, 2021, pp. 11 517–11 529.

[288] S. Zhang, "Quantum strategic game theory," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pp. 39–59, 2012.

[289] 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, 1997, pp. 71–91.

[290] J. Ziv, "On universal quantization," *IEEE Transactions on Information Theory*, vol. 31, no. 3, 1985, pp. 344–347.

[291] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, 1977, pp. 337–343.