Faster Relative Entropy Coding with Greedy Rejection Coding

Gergely Flamich* Stratis Markou*

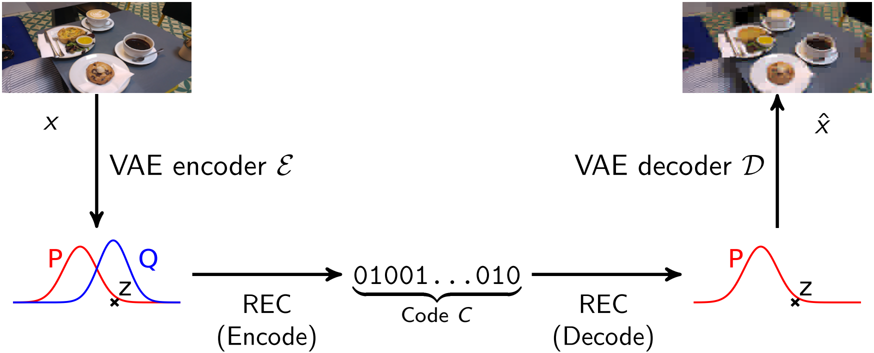
José Miguel Hernández-Lobato

Computational and Biological Learning Lab
Department of Engineering

UNIVERSITY OF CAMBRIDGE

Motivation

Learned compression with VAEs



✓ Does not require quantizing z.

✓ Lossless/lossy compression, private fedrerated learning and others.

Relative Entropy Coding

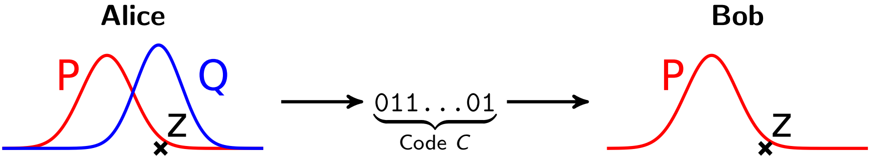
Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

- Is decodable by Bob.
- Represents exact sample from Q.



- As small codelength $|C|$ as possible.
- As short runtime as possible.

Previous work and challenges

Runtime of general REC (Agustsson and Theis, 2020)

Without additional assumptions, any REC scheme will have

$$\Omega(\exp(D_{\mathrm{KL}}[Q\|P]))$$

expected runtime.

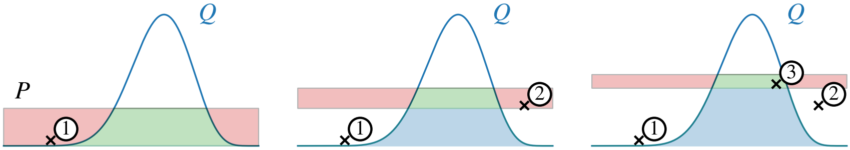
Runtime of A* coding (Flamich et al., 2022)

For 1D, unimodal q/p, the expected runtime of A* coding is

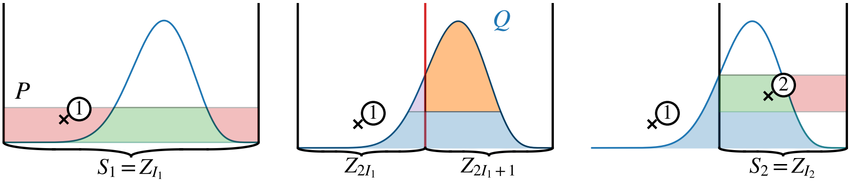
$$\mathbb{E}\left[T\right] = \mathcal{O}(D_\infty[Q\|P]) = \mathcal{O}\left(\log \sup_{z \in \mathbb{R}} \frac{q(z)}{p(z)}\right).$$

Greedy Rejection Coding

We extend a rejection-sampling REC algorithm (Harsha et al., 2007):



Augment algorithm by partitioning the sample space:



Theoretical Results (Informal)

Correctness of GRC

Let q be the target and p the proposal distribution. Let X be the sample returned by GRC. Then, under mild assumptions, GRC terminates with almost surely and $X \sim q$.

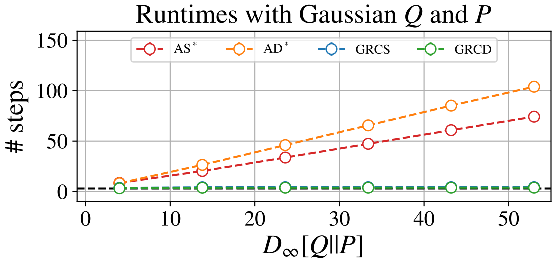
Runtime and codelength of GRC

Let T denote the number of steps GRC takes to run. For 1D unimodal q/p, there exists a partition process, such that

$$\mathbb{E}\left[T\right] = \mathcal{O}(D_{\mathrm{KL}}[Q\|P]).$$

$$\mathbb{H}[X \mid S] \leq D_{\mathrm{KL}}[Q\|P] + 2\log(D_{\mathrm{KL}}[Q\|P] + 1) + \mathcal{O}(1).$$

Experimental Results



Training objective	# latent	Total BPP with ζ coding	Total BPP with δ coding
ELBO	20	1.472 ± 0.004	1.482 ± 0.004
	50	1.511 ± 0.003	1.530 ± 0.003
Modified ELBO	20	$\mathbf{1.470 \pm 0.004}$	1.478 ± 0.004
	50	$\mathbf{1.484 \pm 0.003}$	1.514 ± 0.003

Summary

We develop Greedy Rejection Coding, an optimal REC algorithm.

For more information, find us at
Great Hall & Hall (B1 and B2), poster #1220
Wednesday 13 Dec 8:45-10:45.