Rejection Sampling is Optimal for Relative Entropy Coding

Spencer Hill[1], Fady Alajaji[1], Tamás Linder[1], Gergely Flamich[2]
[1]Queen's University, Canada, [2]Imperial College London, UK

Paper

The minimum achievable rate for relative entropy coding / channel simulation is $I_{\mathrm{F}}(X \to Y)$.

Problem Setup

Stochastic codes

For $X, Y \sim P_{X,Y}$, a *stochastic code* is the triplet $(Z, \mathrm{enc}, \mathrm{dec})$:

- Z – *common randomness*, $Z \perp X$
- $\mathrm{enc} : \mathcal{X} \times \mathcal{Z} \to \{0,1\}^*$ – *encoder*
- $\mathrm{dec} : \{0,1\}^* \times \mathcal{Z} \to \mathcal{Y}$ – *decoder*

satisfying the identity:

$$\mathrm{dec}(\mathrm{enc}(x, Z), Z) \sim P_{Y|X=x}, \tag{1}$$

Minimum achievable rate

For $X, Y \sim P_{X,Y}$, the minimum achievable rate is:

$$\begin{aligned} R^* = \inf_{(Z,\mathrm{enc},\mathrm{dec})} \; & \mathbb{E}[|\mathrm{enc}(X, Z)|] \\ \text{s.t.} \; & \mathrm{dec}(\mathrm{enc}(x, Z), Z)) \sim P_{Y|X=x} \end{aligned} \tag{2}$$

Characterising the minimum achievable rate

Theorem (Li, 2025):

$$I(X;Y) \le R^* \le I(X;Y) + \log(I(X;Y) + 1) + 3.74$$

Question: Can we do better?

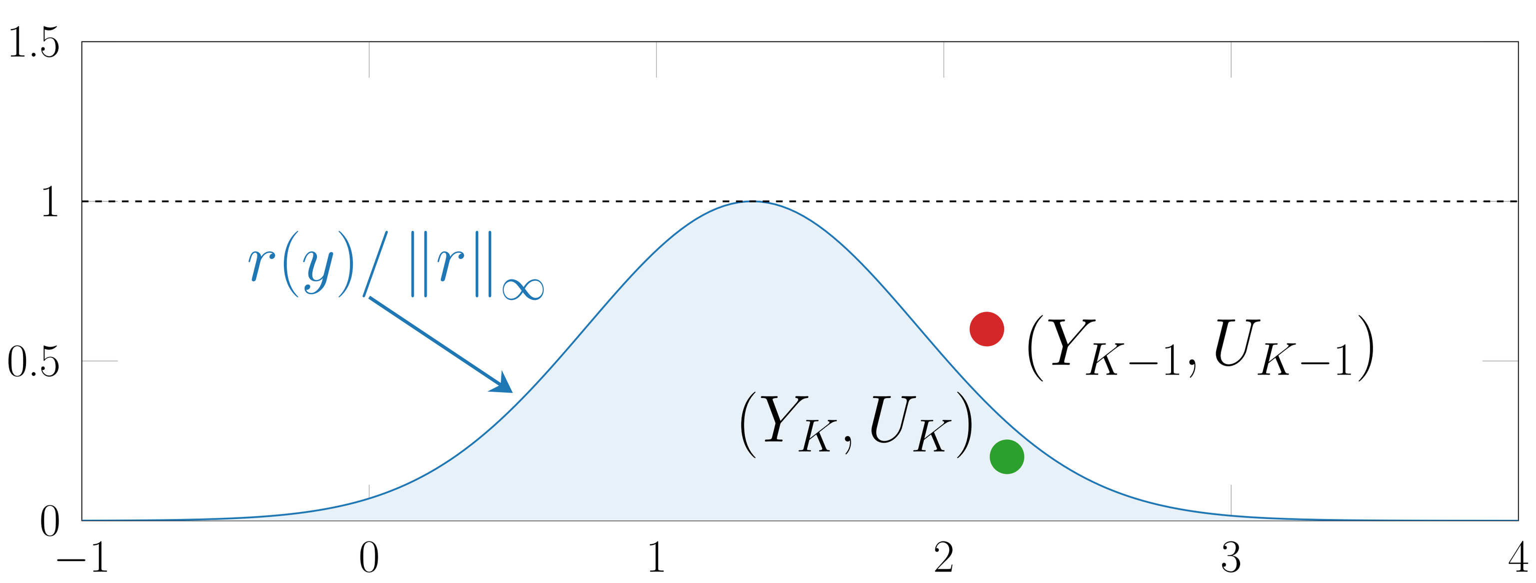
Background

Rejection Sampling

For $\mu \ll \nu$, set $r = d\mu/d\nu$. For $Y_k \sim \nu$ and $U_k \sim \mathrm{Unif}(0,1)$:

$$K = \min\left\{ k \in \mathbb{N} \;\middle|\; U_k \le \frac{r(Y_k)}{\|r\|_\infty} \right\} \tag{7}$$

Then, $Y_K \sim \mu$.



Width function and functional information

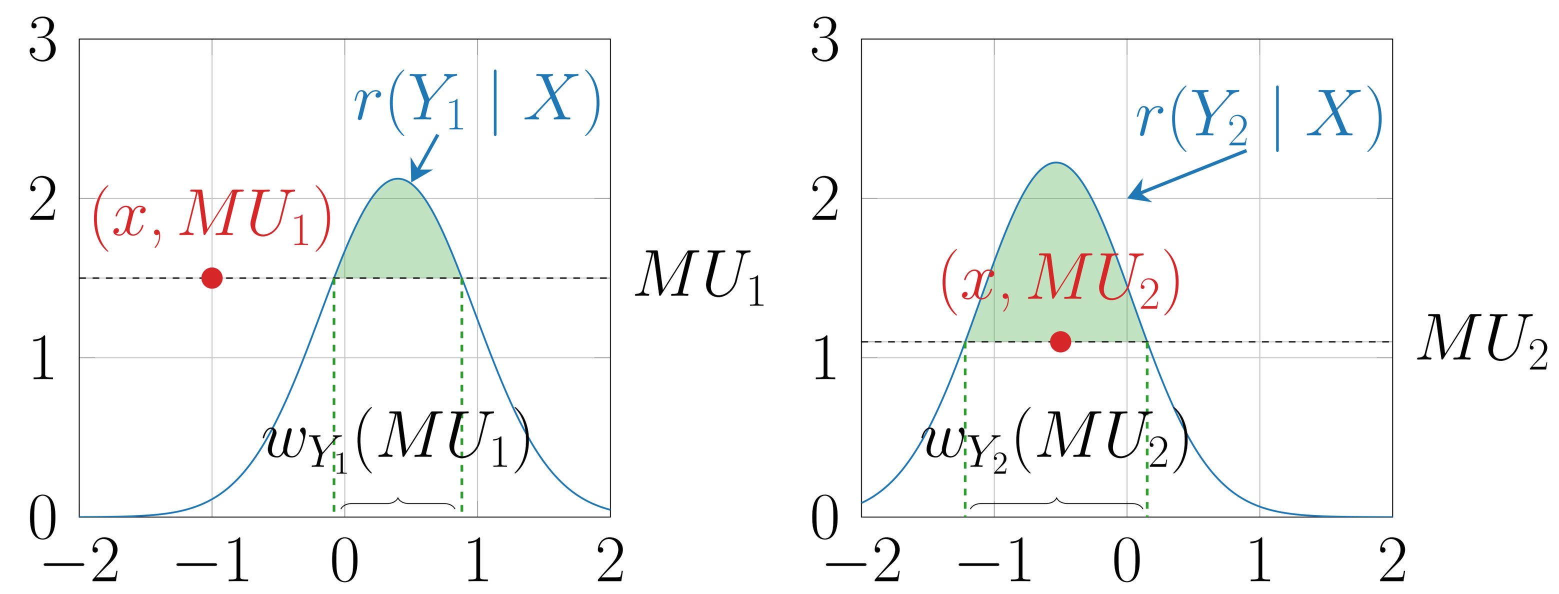
For $X, Y \sim P_{X,Y}$, the *width function* is:

$$w_y(\ell) = \mathbb{P}_{X \sim P_X}\left[\frac{dP_{Y|X}}{dP_Y}(y \mid X) \ge \ell \right] \tag{8}$$

Note: w_y is a PDF. Then, the *functional information* is:

$$I_{\mathrm{F}}(X \to Y) = \mathbb{E}_{Y \sim P_Y}[h[w_Y]] \tag{9}$$

Optimal Compression with the Ring Toss Code



Ring Toss Code

Let $X, Y \sim P_{X,Y}$. Assume $\left\| \frac{dP_{X,Y}}{dP_X \otimes P_Y} \right\|_\infty \le M < \infty$. Then, set $Z = \{(Y_k, U_k)\}_{k=1}^\infty$, where $Y_k \sim P_Y$. Next, set

$$K(x, Z) = \min\left\{ k \in \mathbb{N} \;\middle|\; U_k \le \frac{1}{M}\frac{dP_{Y|X}}{dP_Y}(Y_k \mid x) \right\} \tag{3}$$

Finally, encode K using Huffman code based on

$$Q_{K|Z}(k \mid z) = w_{y_k}(Mu_k) \prod_{i=1}^{k-1} (1 - w_{y_i}(Mu_i)). \tag{4}$$

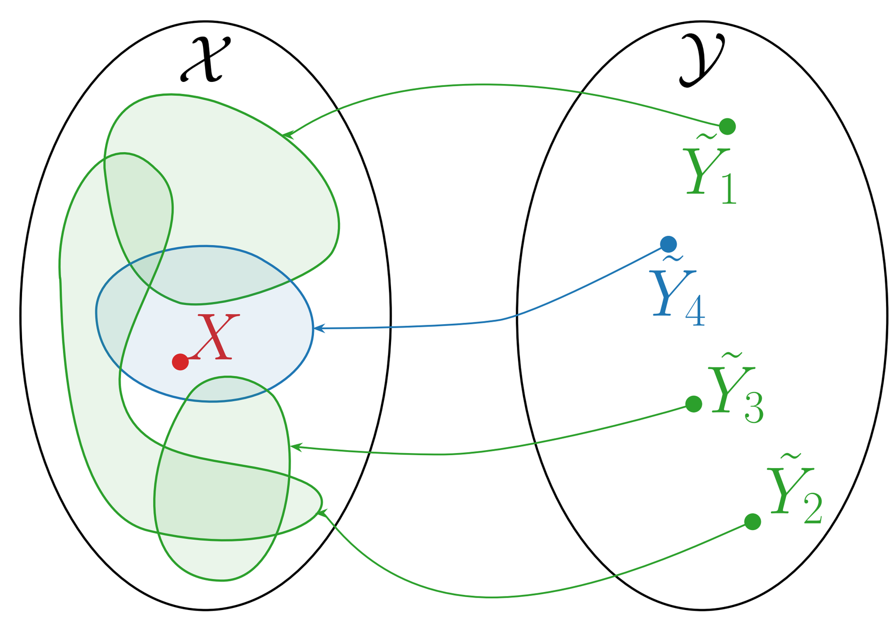
Codelength of RTC

Let $X, Y \sim P_{X,Y}$ and K be the ring toss code for $X \to Y$:

$$H[K \mid Z] \le I_{\mathrm{F}}(X \to Y) + \log e \tag{5}$$

Corollary in conjunction with Flamich et al. (2025):

$$I_{\mathrm{F}}(X \to Y) \le R^* \le I_{\mathrm{F}}(X \to Y) + 2 + \log e \tag{6}$$



Singular Channel

For $X, Y \sim P_{X,Y}$, then $X \to Y$ *singular*, if there exists $g(y)$:

$$\frac{dP_{Y|X}}{dP_Y}(y \mid x) = g(y) \quad P_{X,Y}\text{-a.s.}$$

References

Flamich, G., Sriramu, S. M., and Wagner, A. B. (2025). The redundancy of non-singular channel simulation. In *ISIT 2025*.

Li, C. T. (2025). Discrete layered entropy, conditional compression and a tighter strong functional representation lemma. In *ISIT 2025*.