Greedy Poisson Rejection Sampling

Gergely Flamich[1]

gf332@cam.ac.uk, [1]University of Cambridge

Paper

Greedy Poisson rejection sampling (GPRS) is the first channel simulation / relative entropy coding algorithm whose runtime is optimally fast.

Problem Setup and Motivation

Standard transform coding:

1. Receive data $Y \sim P_Y$
2. Transform the data and quantize it:

$$X \leftarrow \lfloor f(Y) \rceil$$

3. Encode X with P_X using entropy coding

Use relative entropy coding instead!

1. Receive data $Y \sim P_Y$
2. Transform the data and **perturb** it

$$X \leftarrow f(Y) + \epsilon,$$

   e.g. can use $\epsilon \sim \mathcal{N}(0, I)$.

3. Encode X with $P_{X|Y}$ using relative entropy coding

Problem:

Cannot learn f using gradient descent!

Benefits:

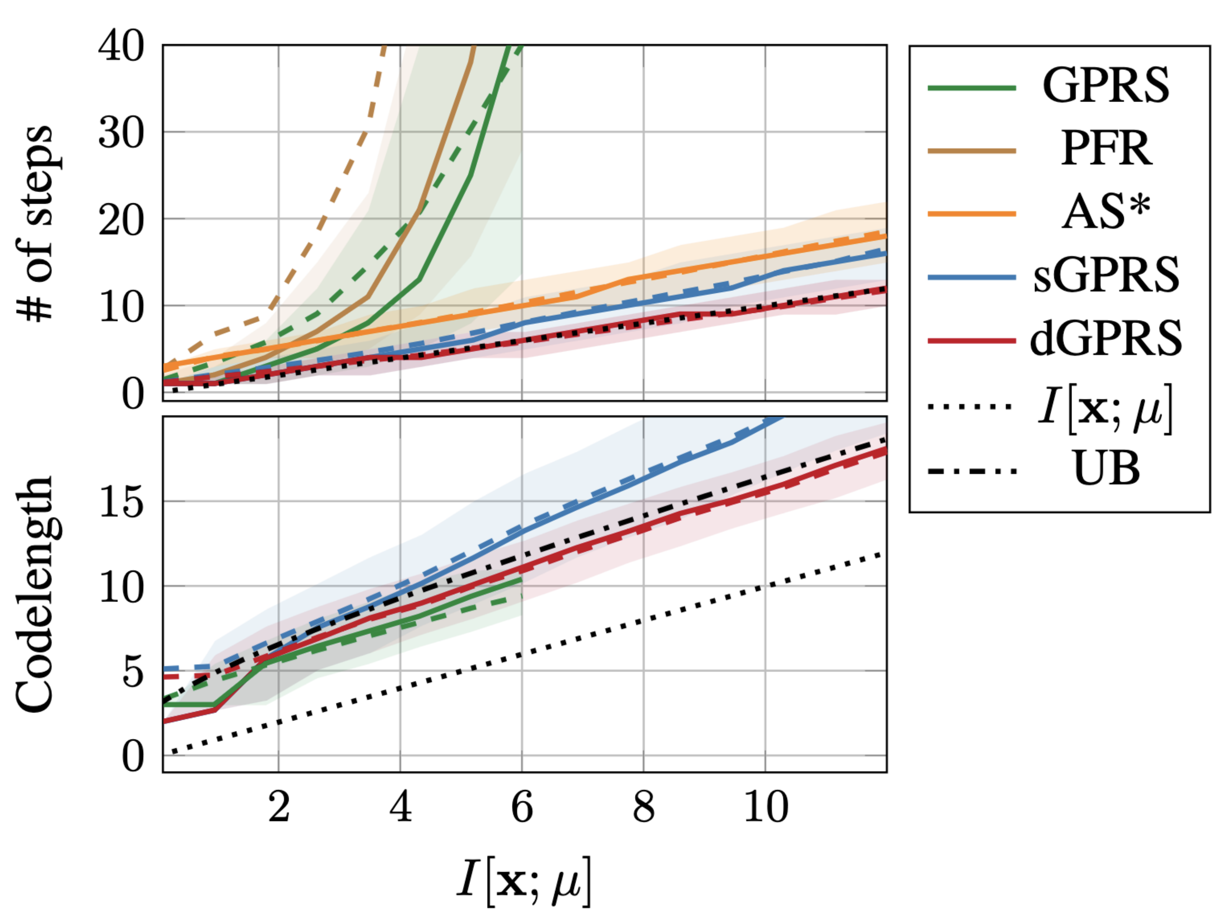
- Can learn f using gradient descent!
- Can incorporate realism constraints!
- Can be used to achieve differential privacy!

Challenge:

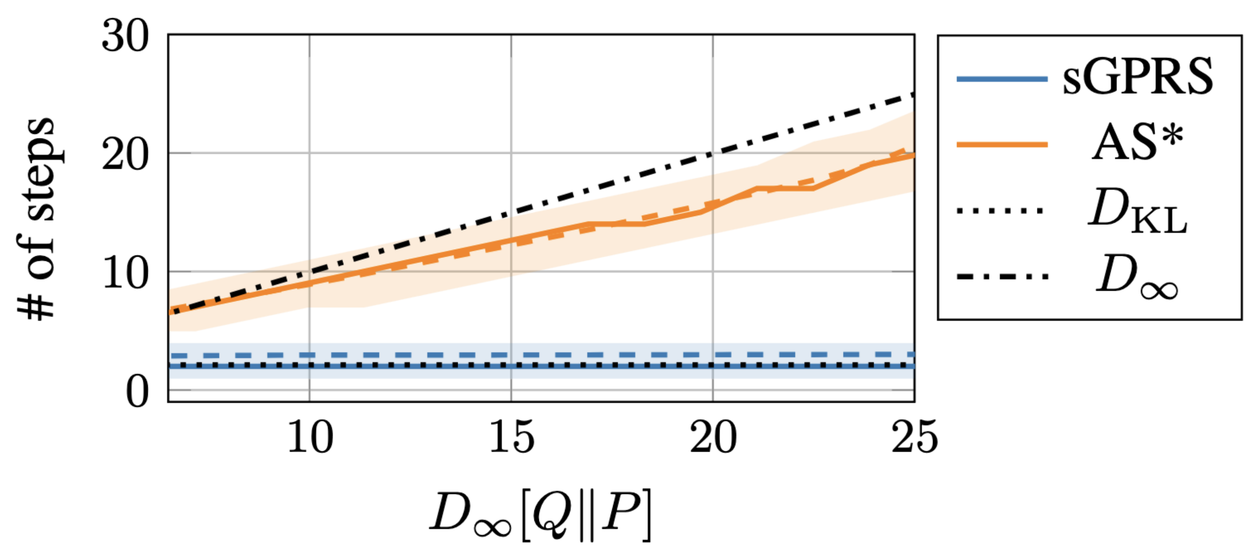
Runtime of general REC (Agustsson and Theis, 2020)

Without further assumptions, any REC scheme has $\Omega(\exp(D_{\mathrm{KL}}[Q\|P]))$ expected runtime.

Comparison with Other Methods

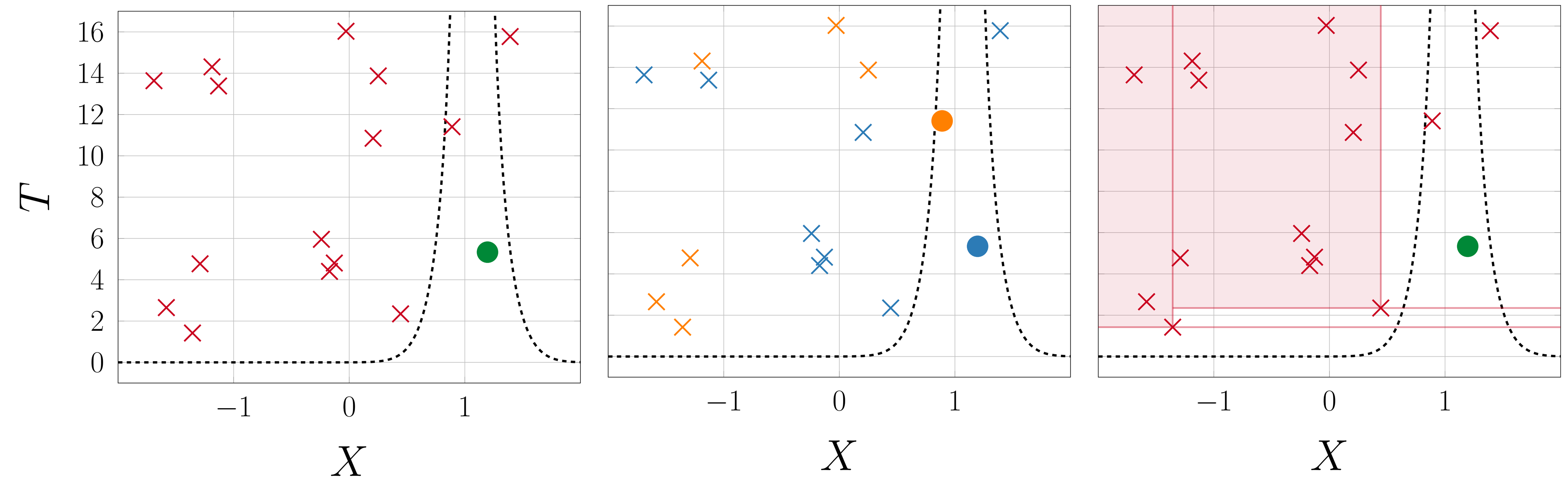


One-shot comparison with A* coding



One-shot runtime scales with $D_{\mathrm{KL}}[Q\|P]$ instead of $D_\infty[Q\|P]$!

Greedy Poisson Rejection Sampling



"standard" GPRS | parallel GPRS | split-on-sample GPRS

High-level idea

- Use Poisson process Π with mean measure $\lambda \times P$, where P is the proposal distribution.
- Accept first point of Π that falls below an appropriately chosen function φ.

How to find φ?

WLOG, assume $\varphi = \sigma \circ \frac{dQ}{dP}$, where Q is the target distribution.

$$\left(\sigma^{-1}\right)' = w_Q\left(\sigma^{-1}\right) - \sigma^{-1} \cdot w_P\left(\sigma^{-1}\right), \quad \text{with} \quad \sigma^{-1}(0) = 0, \tag{1}$$

where $w_P(h) \stackrel{def}{=} \mathbb{P}_{Z\sim P}\left[\frac{dQ}{dP}(Z) \geq h\right]$.

Codelength of channel simulation protocol using GPRS (informal)

Let $X, Y \sim P_{X,Y}$, and C be the code returned by GPRS applied to X, Y. Then

$$\mathbb{E}[|C|] = \mathbb{I}[X : Y] + \log(\mathbb{I}[X : Y] + 1) + \mathcal{O}(1).$$

Runtime of split-on-sample GPRS (informal)

For P, Q over $\mathbb{R}$ with unimodal q/p, the expected runtime D of split-on-sample GPRS is

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