Low-Complexity Locally Private Compression at Finite Blocklengths

Supervisor: Prof. Aslan Tchamkerten
Télécom Paris, Institut Polytechnique de Paris
aslan.tchamkerten@telecom-paris.fr

Context and Motivation

The paper "*A Simple Low-Complexity Locally Private Compression Scheme*" shows that a source can be compressed near its entropy rate while allowing **private local decoding**—any symbol can be recovered by probing only a few compressed bits that reveal no extra information about the other symbols. The current construction, however, relies on a **concatenated block structure** (pad–permute–compress) whose asymptotic guarantees require very large blocklengths. This makes the scheme theoretically elegant but impractical at moderate lengths.

Objective

Design and analyze a **simple, computationally efficient compression scheme** that:

- Achieves rates close to the entropy of the source.
- Supports private local decoding (as defined in the paper).
- **Avoids concatenation**, so that good performance holds at finite blocklengths.

Key Challenges

1. **Privacy without blocks**: Ensuring that the decoder's queries leak no information beyond the requested symbol without relying on independent large blocks.

2. **Complexity bounds**: Keeping both encoding and local decoding polynomial (or near-linear) in sequence length.
3. **Finite-length analysis**: Deriving sharp probability-of-error bounds without asymptotics.

Possible Directions

- Investigate *single-shot* or streaming encoders with local privacy guarantees.
- Explore combinatorial designs (e.g., sparse random graphs, expander-based hash families) to randomize symbol locations instead of block permutations.
- Adapt interactive or universal compression ideas (enumerative coding, polar transforms) to support private local queries.
- Derive information-theoretic limits: minimal number of queried bits for private recovery at finite n.

Expected Outcomes

- A constructive scheme (algorithm + analysis) demonstrating non-concatenated locally private compression with provable rate and complexity guarantees.
- Finite-blocklength bounds (rate vs. error probability).
- A comparative evaluation against the pad–permute–compress baseline.

Prerequisites

- Strong background in information theory (source coding, entropy, finite-blocklength analysis).
- Familiarity with randomized algorithms and basic probability.
- Coding skills (Python/Julia/Matlab) for experimentation.

Deliverables

- Mid-term report: survey of existing locally decodable/private compression techniques and proposed approach.
- Final report: theoretical analysis, algorithm, and numerical experiments.
- Optional: implementation of a prototype encoder/decoder.