EE388 - Modern Coding Theory

As a final requirement, you are asked to get involved in a project:

1. Form a small team (ideally 2 students).

2. Choose a research topic, and submit to Andrea your team composition and the topic by Sunday May 6.

3. Review the relevant literature: note that the references provided here are only entrypoint and not an exhaustive bibliography.

4. Think independently of the problem. Depending on the case, this might mean: implement and simulate some of the solutions in the literature; try to improve existing analysis; propose alternative solutions (or problem formulations).

5. Give a presentation in class about your work. The presentation should be accessible to all the other students. Attendance is required.

Here is a list of possible topics with some description and references. More details (and more suggestions) are available from Andrea. You are also free to suggest alternative topics.

Note that the outlines below should provide a baseline for your work, but you are welcome to do more!

**Analysis of spatially coupled codes.** The objective of this project is to study and present the proof that spatially coupled codes achieve capacity over general BMS channels. In particular:

- Study proofs in the literature. A few references are [KRU11, KRU13, YJNP12], but you should look for other relevant papers.
- Present a self-contained outline of the proof (I would suggest to focus on the erasure channel).
- Carry out some numerical simulations to demonstrate the practical behavior of this code.
- A plus would be to think of ways to improve over existing proofs.

**Coding-theoretic ideas in compressed sensing.** Ideas from coding theory have been used to propose specific constructions of sensing matrices in compressed sensing. Examples include [Abb10, HAT12] (which use polar codes) and [KP10, KMS+12, DJM13] (which use spatially coupled matrices). The objective of this project is to survey these and related constructions. You should

- Provide necessary background on this problem.
- Overview these constructions, focusing on one.
- Overview results for these constructions and proof ideas
- Carry out numerical simulations to demonstrate their behavior.
A plus would be to think about approaches to improve over these constructions.

**Error floor properties of LDPC codes.** A major concern with LDPC codes is that randomly generated graphs can show poor behavior below threshold because of small subgraphs that cause localized errors. This phenomenon is well understood for the erasure channel, since there is a good understanding of what subgraphs cause errors [DPT02]. In the case of general BMS channels, there is no canonical choice of the ‘bad subgraphs.’ An empirical approach based on importance sampling is described in [Ric03].

- Overview the general problem of finite length analysis and error floors.
- Overview results in the literature.
- Provide more detail about the one of following
  1. Binary erasure channel. In this case explain the analysis of the typical size of stopping sets for LDPC ensembles, as initiated in [DPT02].
  2. General BMS. In this case explain and implement the importance sampling method of [Ric03].

**Making polar codes practical.** In practice, polar codes present a gap to capacity that decreases quite slowly with the blocklength. Several ideas have been proposed to overcome this problem, including strengthening the code and improving the decoding algorithm. A couple of entrypoints in this literature are provided by [MHU14, TV15], but there are many other relevant papers you should search for.

- Present the problem.
- Discuss the most important ideas to overcome it, and why they are expected to work.
- Show in simulation how they perform.
- Think of possible ideas to obtain further improvement.

**Codes for lossy data compression.** Often a good error correcting code can be turned into a good source code for lossy data compression. Channel capacity is replaced by the rate distortion curve. Examples of this type are provided by [WM05] (using LDPC codes), [KU10] (using polar codes), [AMUV12] (using spatially coupled ensembles).

- Present the lossy source coding problem, and its fundamental limits.
- Present in detail at least one of the approaches using good error correcting codes.
- Survey other approaches.
- Show in simulation how your favorite method performs.

**Message passing algorithms for random constraint satisfaction.** Message passing algorithms can be used as heuristics to perform inference in graphical models, and to solve random constraint satisfaction
problems. Notable example of the last class include $k$-satisfiability and graph coloring. The objective of this project is to understand how these algorithms (that we studied in the context of coding) can be applied to very different problems. A few references include [MPZ02, MMW07, MM09, MRTS07] (the simplest reference is probably the last one).

- Describe the problem.
- There exist several variants of the message passing approach. Describe and explain precisely at least one of these variants.
- Describe what is the expected behavior of this algorithm for random instances (what is known, what is not known, what is conjectured).
- Implement at least one of these algorithms and test it on randomly generated instances.

References


