Yuanxin Guo Homepage: https://guoyuanxinkevin.github.io/ E-mail: yuanxinguo@link.cuhk.edu.cn

Statement of Objectives

My research interest lies in information theory and coding theory, a subject at the crossroads of mathematics, computer science, and electrical engineering. A recurring pattern appears in several of my research experiences: I enjoy using mathematics of all kinds to solve practical problems. As a math major, I obtained a solid mathematical background by maintaining a good GPA (3.9/4.0), taking graduate-level mathematics courses, and holding a few teaching assistant positions. Meanwhile, I have also gained sufficient exposure to research by taking a research position at various institutes and laboratories. Out of my enthusiasm for research, I decide to apply for a Ph.D. in EECS at MIT, allowing me to pursue my interests further and work with top-tier scientists in the world.

My first encounter with research was in the Summer of 2019. I joined the Network Coding Lab at the Chinese University of Hong Kong, Shenzhen, under the supervision of Prof. Shenghao Yang and Dr. Ximing Fu. My first project was about shift-XOR codes, a novel family of erasure-correcting storage codes with low computational complexity. One of the challenges is to design generator matrices that entail low storage overheads. I started the literature review from our lab's paper which introduced refined increasing difference (RID) matrices and showed by calculation that optimal RID matrices outperform other generator matrices in terms of storage overhead. In addition, I designed a 'puncturing' technique for non-systematic RID shift-XOR codes that reduces the storage overhead of an (n, k) code by a fraction of k/n. However, the practicality of non-systematic codes is limited, and the contribution of this work is marginal. The manuscript I had written eventually failed to be published. Nevertheless, I did not give up exploring. During the second year at the lab, our group designed a generalization of RID matrices called two-tone matrices. Using combinatorics, I proved that a variant of Vandermonde matrices are optimal two-tone matrices as they incur the lowest storage overheads. Furthermore, I showed that the extra storage cost of the optimal two-tone code is less than 50% of that of the optimal RID code. This work eventually led to a co-authored conference paper at ISIT 2021 and a published China patent.

The success of designing an applicable code aroused my interest in coding theory. More importantly, it stimulated my ambition of doing something more fundamental. Therefore, when Prof. Kenneth W. Shum and Prof. Shenghao Yang asked me if I wanted to develop a general theory of shift-XOR codes, I nodded without any hesitation. Thanks to my mathematical training, I could formulate ideas and generalize concepts using abstract and rigorous language. Instead of being restricted to coding theory, our subject became infinite shift-add systems whose symbols take values from an arbitrary abelian group. I observed that solving a variable in such a system is equivalent to reducing it to another such system. Inspired by a graph theory course, I proposed representing these systems as vertices and connecting them via directed edges. It turns out that this graph-theoretical characterization is closely related to a simple decoding algorithm called successive zigzag decoding. In a conference paper that I co-authored on the subject, we have given a necessary and sufficient condition on a shift-add system for such an algorithm to be applicable. Surprisingly, the nontrivial coding theoretical result relies on several simple graph-theoretical lemmas proven by me. I presented this research as an oral presentation at ISIT 2021, yet what I have gained from this experience exceeded a paper by far: A different formalism or tool might reshape the same problem entirely; I should choose the one that fits my purpose. Currently, a journal paper on the subject is in preparation. The year-long research on shift-XOR codes had walked me into the gate of coding theory, and I willingly decide to take a step further. At the beginning of 2021, I joined the Wireless Technology Lab of Huawei Technologies. During the internship, I acquired vast stores of knowledge on coding theory and information theory, especially on polar codes, Reed-Muller codes, and non-asymptotic information theory. I also witnessed unfamiliar parts of mathematics, say concentration inequalities and group actions, come into play in my field of interest. Extending our group's previous work on improving polar codes, I managed to establish the equivalence between pre-transformed and parity-check polar codes using linear algebra. Furthermore, I explicitly constructed a pre-transformation for polar codes that could improve the distance spectrum. This work provides substantial guidance on

code design, and a first-author paper on this work was accepted and presented as poster at ISTC 2021.

Outside of coding theory and information theory, I have made plenty of attempts to explore and learn new things. In the summer of 2021, I joined Prof. Stark Draper at the University of Toronto as a summer research intern. We sought to recover the underlying community structure of a network in the presence of unknown geometric information, modeled by a geometric block model (GBM). After studying a state-of-the-art recovery algorithm based on semidefinite programming (SDP) for stochastic block models (SBM) and a motif-counting algorithm for GBM, I adjusted the SDP algorithm to make it better fit the community detection task of GBM. Experiments had shown that my improved algorithm has a much higher success rate compared to the original SDP algorithm and the naïve motif-counting algorithm (99% compared to 3% and 0% in one instance). Prior to that, I worked with Prof. Zhi-Quan Luo at the Shenzhen Research Institute of Big Data on a project on intelligent reflecting surfaces (IRS). In contrast to a traditional wireless communications approach to optimize the configuration of an IRS, we followed a statistical path. Inspired by the RFocus algorithm proposed by CSAIL, I first extended the algorithm to a more general setting; then I designed a series of new algorithms using similar idea. Using the extended RFocus as baseline, all of the new algorithms yielded better received signal strength indicator (RSSI) in numerical simulations. These experiences opened up different possibilities to me, and helped me to pinpoint my interest eventually.

Besides research, another zeal of mine is sharing knowledge with more people. From 2019 to 2021, I have been an undergraduate teaching fellow (USTF) for four mathematics courses, including Calculus I, Complex Variables, Linear Algebra, and Abstract Algebra. For the Abstract Algebra course, I have even typeset extended reading materials and practice questions for students. Due to my passion in teaching, I was rated 5.94 out of 6 in an end-of-term teaching evaluation. Additionally, I engaged actively in independent study groups for three terms, where I gave tutorials on topics in Numerical Linear Algebra, Markov Chain Monte-Carlo, and Actor-Critic Reinforcement Learning. My supervisors, Prof. Zhi-Quan Luo and Prof. Jim Dai, had both given me an A-grade. Currently, I am co-organizing an Applied Category Theory seminar with Prof. Kenneth Shum. These experiences has fostered my presentation skills and leadership, which are equally important as research capability for a successful Ph.D.

Personally, I am determined to devote my entire life to research and I aim for obtaining a tenure-track faculty position at a university. I believe that obtaining a Ph.D. in Information Theory and Coding at MIT would bring me one step closer to my lifetime goal. As a world-renowned institute, MIT is known for its strength in engineering subjects and its world-class faculty and students. I am particularly excited to work with Prof. Yury Polyanskiy. Impressed by his ground-breaking Ph.D. thesis on finite-blocklength information theory, I become more overwhelmed each time as I discovered the diversity of topics he touches on. At this point, his work in information-theoretic methods in statistics fascinates me the most, and I am willing to learn and contribute. If I am given the opportunity to attend graduate school at MIT, I am confident that I can conduct cutting-edge research and set off my academic career.