GENERAL ARITHMETIC CODING FOR QUANTUM COMPRESSION

Cong Yu, Todd Brun

Department of Electrical and Computer Engineering
University of Southern California

Abstract

Based on an existing algorithm for reversible quantum arithmetic compression of a quantum Bernoulli source, a generalized compression procedure is proposed for quantum sources of any dimension. This hand compilation of classical arithmetic compression can compress quantum source distributions more complex than their classical counterparts. By this procedure the compression rate approaches the entropy of the quantum source asymptotically while the expected fidelity approaches 1 asymptotically as well.

Introduction

Quantum compression has been regarded as a largely theoretical (rather than practical) process. However, it could play an important role in quantum information processing, just as the classical compression does. Therefore it is important to find algorithms for practical quantum compression; and this is challenging, because most classical compression algorithms do not adapt well to the quantum case. Asymptotically, the counterpart of classical Shannon compression is Schumacher's quantum compression[4, 2], which has proven that the theoretical lower bound of any viable compression scheme is the Von Neumann entropy of the quantum source's density operator, regardless of the dimensions of the qudits. We give a practical procedure to compress general quantum strings.

Methodology

Classical Arithmetic Coding

The goal of classical block compression is to construct a function that maps the original random variable with some known distribution to a new random variable with uniform distribution. The new variable is incompressible and forms the compressed version of the original information. Arithmetic coding, motivated by Shannon-Fano-Elias coding, can compress any sequence of I.I.D. elements at a rate slightly higher than the asymptotic limit of the Shannon entropy. Classical arithmetic coding maps a sequence to a subinterval of $[0, 1)$, and the midpoint of the interval represents the codeword.

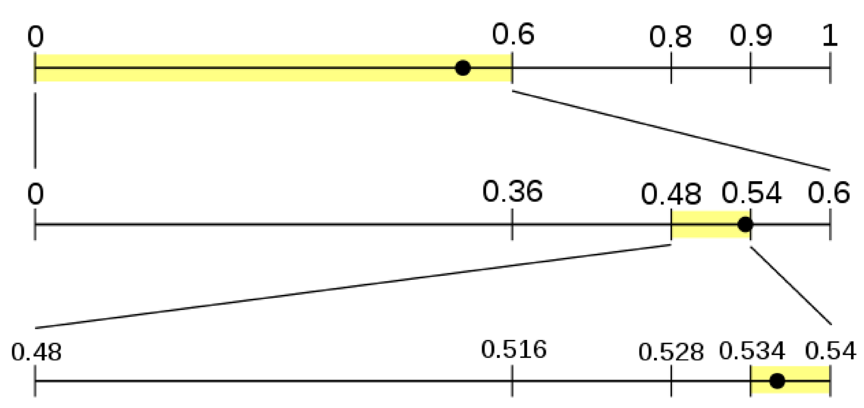

Figure 1: The decoding of 0.538 (the circular point) in the example model. The interval is divided into subregions of length proportional to the symbol frequencies, then successively subdivided in the same way[5]

Scheme of the Quantum Encoder "E"

This circuit encodes the sequence $|\chi\rangle$ from the quantum source to a compressed string $|C_{\mathrm{mid}}\rangle$, represented as a fixed-point register. This circuit uses submodules "e" and "d" described below. The decoding block "D" is the inverse of this circuit. These blocks play the same role as the circuits in [1].

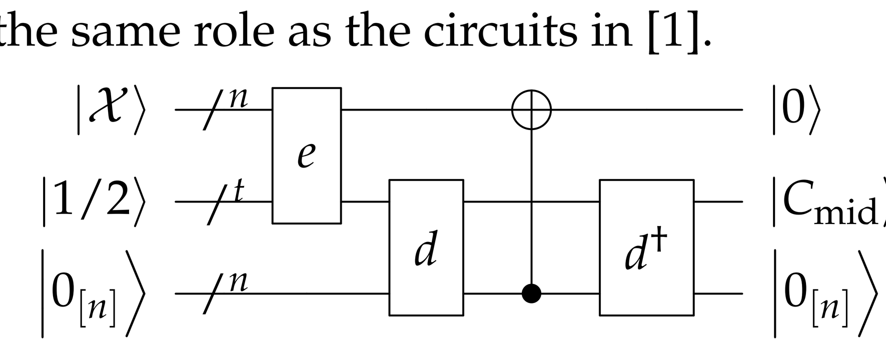

Figure 2: Quantum circuit to remove entanglement between the codeword and input sequence. Modules described below.

The "e" Module

The input sequence is $|\chi\rangle = |X_1\rangle \otimes \cdots |X_n\rangle$ where $X_1, \ldots, X_n$ are I.I.D. random variables. Instead of iterating the upper and lower bounds of the interval that represents the sequence χ, we use the midpoint and length of the interval. This changes the iteration from the classical case by a straightforward calculation. The length represents the probability of the sequence. It can also absorb typical measurement block described below.

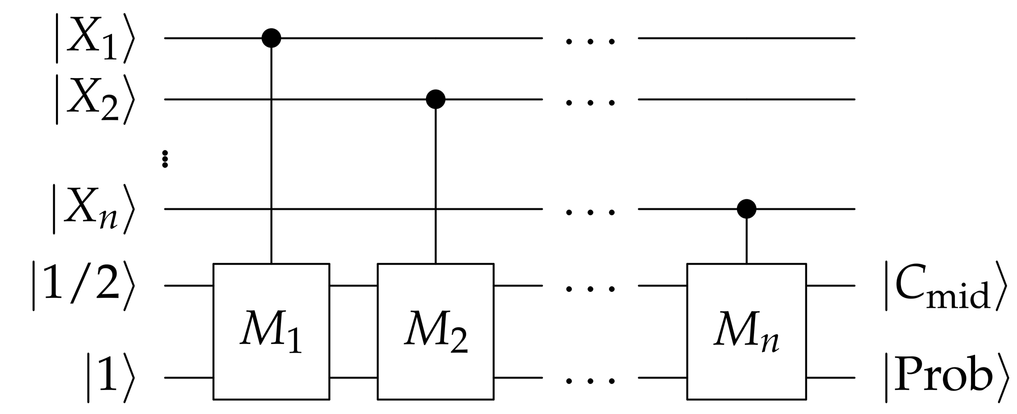

Figure 3: Quantum circuit for the "e" module. M_i encodes the symbol X_i and iterates the values of C_{mid} and Prob.

Strong Typical Measurement

A major difference between classical and quantum compression is that a fixed classical compression procedure cannot compress every sequence—in particular, the atypical ones—but the probability of atypical sequences is small. Hence on average the compression can be achieved. However, a quantum sequence could be a superposition of all basis vectors, including typical and atypical sequences. We have to compress the sequences in the typical subspace and discard the atypical part. This circuit block can be absorbed into the "e" module.

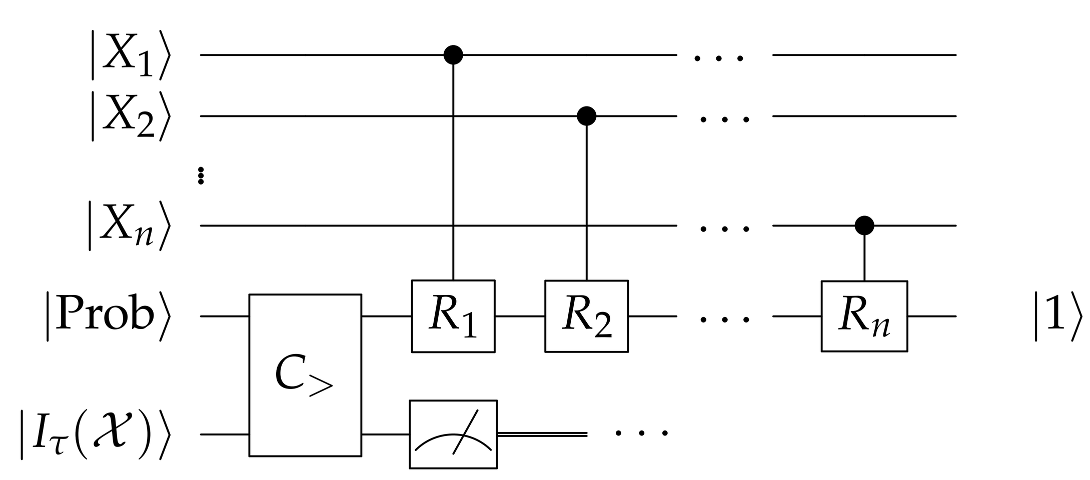

Figure 4: Quantum circuit to measure strong typicality and reset $|\mathrm{Prob}\rangle$ to $|1\rangle$.

The "d" Module

We must remove any entanglement between the midpoint variable and the input sequence after "e" module. Borrowing an idea from [3], we transform the classical decoding procedure to a quantum version. This lets us do the entanglement removal.

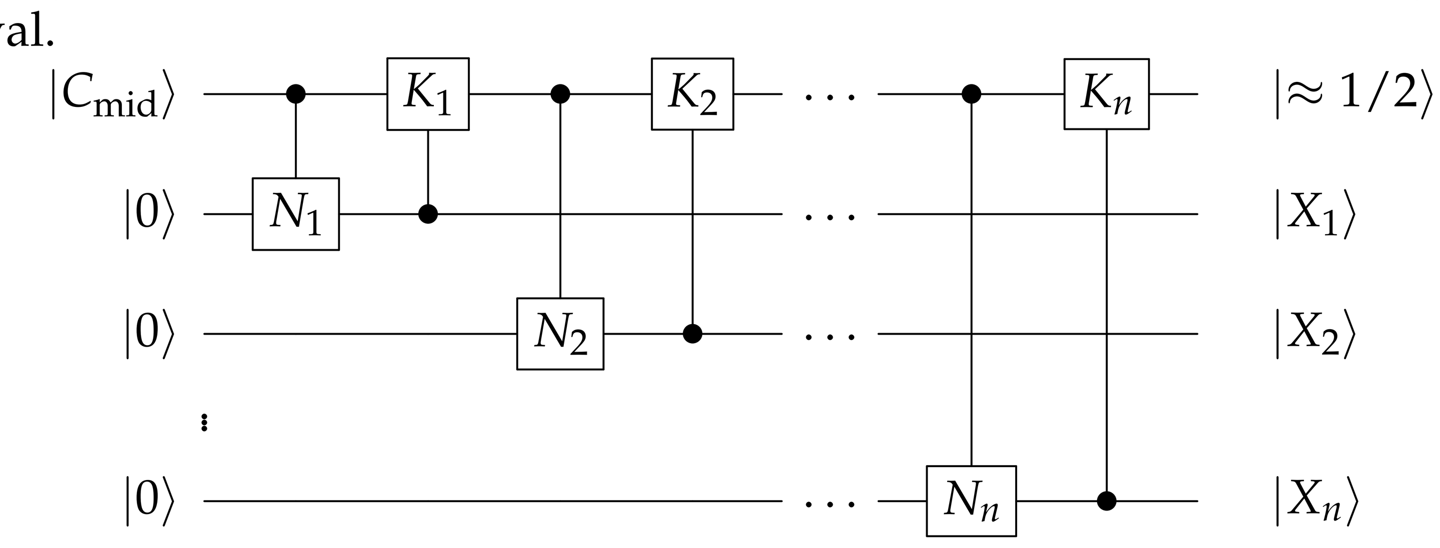

Figure 5: Quantum circuit for the "d" module. N_i extracts the ith symbol from C_{mid}, while K_i iterates the value of C_{mid} itself during the decoding procedure. The final value of C_{mid} may not be exactly $1/2$ because of finite precision arithmetic

Then we already have all the elements [1] to truncate the midpoint variable.

References

[1] I. L. Chuang and D. S. Modha. Reversible arithmetic coding for quantum data compression. *IEEE Transactions on Information Theory*, 46(3):1104–1116, 2000.

[2] R. Jozsa and B. Schumacher. A new proof of the quantum noiseless coding theorem. *Journal of Modern Optics*, 41(12):2343–2349, 1994.

[3] J. Langford. Generic quantum block compression. *Physical Review A*, 65(5):052312, 2002.

[4] B. Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995.

[5] Wikipedia contributors. Arithmetic coding — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Arithmetic_coding&oldid=929412375`, 2019. [Online; accessed 24-December-2019].