Quantumized Graph Cuts in Portfolio Construction and Asset Selection

W. Bernard Lee*
HedgeSPA Limited, 29 Mediapolis Circle #04-14, Singapore
School of Accounting and Finance
Hong Kong Polytechnic University, Hong Kong
bernard.lee@hedgespa.com
bernardl@cs.stanford.edu

Anthony G. Constantinides
Financial Signals Processing Laboratory
Imperial College London
Exhibition Road
London SW7 2AZ, United Kingdom
a.constantinides@imperial.ac.uk

ABSTRACT

This paper is concerned with two fundamental problems in investment science, namely a) the construction of a portfolio by segregating target assets into sectors representing a typical market index and b) the selection of target assets from each of those sectors. Such solutions may be applied, for example, to construct a portfolio of 50 assets (say) that aims to outperform the S&P 500 index by selecting the most promising performer in each applicable sector or subsector. The formulation of this investment objective is non-trivial due to the following reasons:

- For Problem 1, the cohorts of assets should be selected based on not only a fundamental classification such as the Global Industry Classification Standard (GICS)[5] but also appropriate statistical characteristics of the assets. The classic argument is that Tesla should be classified as a Technology firm instead of an Automotive manufacturer.
- For Problem 2, the selection of assets with each sector can be driven by factors that may be non-uniform across different sectors. The classic example here is that Inventory is usually not important in the Software Technology sector but it is considered important for the Industry Manufacturing sector.

A computationally efficient approach to solve a similar problem using Graph Theory has been described in Chapter 13 of *Data Analytics on Graphs*[4], contributed by a co-author of this paper. Problem 1 can be solved in the Graph Theory sense by cutting the Set A of assets into two disjoint subsets to maximize the difference in a certain performance metric $f(x_1, x_2, \ldots, x_n)$ of one subset $\{x_1, x_2, \ldots, x_n\}$, where $n < N$, to the same metric applied to the remaining subset of $N-n$ elements. This problem, known as finding the "maximum graph cut" or Max-Cut, sounds intuitively straightforward for a small number of exemplar assets, but its complexity can be prohibitive when applied to real-world, industrial-scale problems. To illustrate the complexity of the problem, if $N = 500$ (e.g. the S&P 500 stock components), the number of combinations to split the vertices into two subsets is $C = 1.6 \times 10^{150}$, and we have to continuously subdivide the leaves in the tree of portfolio cuts to arrive at a suitable sector partitioning that combines fundamental classifications with asset statistical characteristics.

Once we have arrived at the suitable sectors (currently there are 11 sectors for the S&P 500 index, resulting in the average of 45 assets in each sector, but the computer may choose say 20 sectors with 25 assets each on average), in theory, we can apply a similar methodology to the factor sets to fit an asset selection model, so that the model (usually via using a neural network) is only allowed to choose one key factor from each subset of similar factors as a workaround to the well-known problem of overfitting models with too many similar factors, resulting in low *ex-ante* predictive powers. Using Graph Theory to enhance neural networks is a technique known as Graph Neural Networks, which has successfully addressed challenging missing data issues for cross-sectional and panel-data models. However, if there are say 40 factors in a single time slice, we will be working with an average of 1000 factors or $N = 1000$ for a single cross-section in time, before even introducing the time dimension to the problem, which may grow each factor set to a few thousand factors.

Given the computational complexities shown above, these problems are not tractable within a meaningful timeframe using classical computers. In a paper presented at the 2021 Annual Meeting of the American Statistical Association[3], the authors have proposed an alternative approach that would allow novel computational means using the quantum computing variant of the Max-Cut algorithm. Following the work done by quantum physicists at MIT[2], we propose to transform the Max-Cut problem that can be solved by "brute force" combinatorial optimization into its equivalent representation using the Ising Hamiltonian function. Instead of millions of CPU cycles to create a single vector of random samples, a quantum computer is designed to generate a vector of simultaneous random samples in a single cycle. This paper aims to describe the initial results of a test implementation from solving this problem on:

- Classical computer as the base performance benchmark,
- Accelerated classical architectures, and
- IBM quantum computers and/or simulators[6].

CCS CONCEPTS

• **Mathematics of Computing**; • **Discrete Mathematics**; • **Graph Theory**; • **Graph algorithms**;

KEYWORDS

Portfolio Cuts, Graph Theory, Time Series Estimates of Expected Returns, Combinatoric Optimization and Simulations, Quantum Computing, Maximum Graph Cuts

*Corresponding author.

Authors' addresses: W. Bernard Lee, HedgeSPA Limited, 29 Mediapolis Circle #04-14, Singapore; School of Accounting and Finance; Hong Kong Polytechnic University, Hong Kong, bernard.lee@hedgespa.com, bernardl@cs.stanford.edu; Anthony G. Constantinides, Financial Signals Processing Laboratory, Imperial College London; Exhibition Road, London SW7 2AZ, United Kingdom, a.constantinides@imperial.ac.uk.