Invited: Challenges and Opportunities of Quantum Optimization in Finance

Zichang He, Shouvanik Chakrabarti, Dylan Herman, Niraj Kumar, Changhao Li, Pierre Minssen, Pradeep Niroula, Ruslan Shaydulin, Yue Sun, Shree Hari Sureshbabu, Romina Yalovetzky, and Marco Pistoia
Global Technology Applied Research, JPMorgan Chase, New York, NY 10017 USA

ABSTRACT

In recent years, immense progress have been made in quantum optimization techniques. In this position paper, we share our opinion on some of the challenges facing the quantum optimization community and highlight opportunities which we believe would benefit from increased attention of researchers.

1 INTRODUCTION

Quantum computers have the potential to deliver broadly applicable speedups in optimization. The ubiquity of hard optimization problems in science and industry underscores the pressing necessity for developing novel approaches to optimization. Financial industry is among the domains in need of better optimization tools, as hard optimization problem underlie many aspects of market functioning and investment decision-making that sustain our economy [6]. Recent years saw rapid progress in quantum algorithms for optimization, with many new techniques introduced and previous techniques improved.

At the same time, quantum algorithms for optimization have not yet outperformed their state-of-the-art classical counterparts in practice. A number of challenges must be addressed before reaching this milestone. In this paper, we present our perspective on the most pressing challenges, and outline some promising opportunities for future advancement. This paper does not aim to provide an exhaustive survey of quantum optimization, and many exciting recent developments fall outside of our scope. The topics we discuss are multifaceted and subject to active research. In the interest of space, we omit references to works that did not originate in our group and occasionally distill years of complex research into a short sentence. Our goal throughout this paper is clarity; consequently, we occasionally exclude caveats and overlook subtleties. Finally, we remark that all the statements we make focus on optimization domain only. The same classes of algorithms may provide speedups in other domains, e.g. in quantum simulation.

Throughout this paper we use the notion of "quantum computational speedup" or simply "speedup". By this we mean that on a fixed problem and for a fixed desired solution quality, a quantum algorithm finds a solution faster than a classical one. We note that obtaining the same solution faster can always be converted to obtaining a higher solution quality within the same time budget.

DAC '24, June 23–27, 2024, San Francisco, CA, USA

ACM ISBN 979-8-4007-0601-1/24/06…$15.00
https://doi.org/10.1145/3649329.3664193

2 CHALLENGES

The challenges we discuss in this section imply that a large proportion of commonly considered quantum techniques for optimization may not become practical in the near future. While these challenges may seem daunting, we believe that acknowledging and tackling them head-on is required for progress.

2.1 Provable speedups are small

The primary challenge in quantum optimization is that the rigorous end-to-end quantum speedups that have been achieved in the past are typically low-degree, often quadratic or worse. These speedups are commonly obtained by applying amplitude amplification to specific components of classical algorithms. The dearth of large provable speedups in optimization motivates the development of a veritable kaleidoscope of heuristics, whose the challenges we discuss below. While super-polynomial proven speedups in optimizations do exist, they are exclusively derived within contrived settings that are unlikely to correspond to realistic industrially-relevant problems. For example, a common technique is to convert the hidden subgroup problem into an optimization problem and use Shor's algorithm to solve it, arriving at a super-polynomial separation between quantum and best classical methods. Despite their mathematical sophistication, such analyses fail to offer insights into whether quantum algorithms can offer super-polynomial speedups for any instances of industrially-relevant problems.

The existing provable low-degree speedups face several challenges. In many cases the speedup is demonstrated over the worst-case instance of a problem and may not be realizable on average-case instances that are actually encountered in industrial settings. Even when the speedup is universal [2], the high overhead of error correction implies that these speedups are unlikely to be realized except for exceedingly large problems that require months of classical computation time. Given that most industrial processes necessitate problem solution within hours at most, these speedups are unlikely to yield significant economic impact.

2.2 Variational algorithms are unlikely to yield speedups

By far, the most popular class of heuristics for optimization consists of variational quantum algorithms or VQAs. VQAs are a family of algorithms that, per problem instance, optimize the parameters in a parameterized quantum circuit to prepare states that minimize specific target objective. However, there are several major barriers to developing useful VQAs.

The first is that, the cost of obtaining the gradient of a parameterized quantum circuit, in terms of circuit executions, scales linearly with the number of parameters. This is in stark contrast to the backpropagation algorithm used by classical neural networks, where the cost, in terms of evaluations of the neural network, of obtaining the gradient remain independent of the number of parameters. While the number of parameters can be reduced in some cases without hurting convergence [15], it remains uncertain if this is feasible in general and whether the magnitude of the reduction is sufficient to meaningfully alleviate the cost of training.

The second is the well-known barren plateau (BP), characterized by an effectively flat landscape that occurs due to gradients exponentially (with the number of qubits) concentrating to zero over random, usually uniform, parameter initializations. This phenomenon occurs in noiseless circuits, but is stronger in noisy cases. Since both the cost function and its gradients must be estimated by measuring the quantum state, this implies that both zero- and first-order optimization methods will require exponentially many samples from the quantum circuit to determine how to make progress in the cost landscape. Recent techniques proposed for overcoming this problem [3] typically produce parameterized circuits that are classically simulatable.

The third issue is the prevalence of spurious local optima that can be significantly deviate from the global optimum. In general, this generally leads to exponentially many parameters being required to ensure quality solutions, making the training infeasible due to the high cost of obtaining gradients.

2.3 Large quantum annealing devices yet to yield speedups

Quantum annealers were the first type of quantum devices capable of tackling optimization problems beyond trivial sizes. Their scalability advantage stems from the fact that the analog control of qubits is easier to engineer than high-fidelity digital operation. This led to devices with thousands of qubits available commercially, offering an enticing opportunity to benchmark such devices against classical solvers on problems for which classical solvers exhibit non-trivial runtime.

Unfortunately, despite over a decade of research, conclusive evidence demonstrating a quantum speedup for optimization problems using quantum annealers remains unclear. For example, recent work [1] has benchmarked instances of the maximum-independent-set problem, previously reported to be solved with a potential super-linear quantum speedup over classical simulated annealing, using a quantum annealer based on two-dimensional Rydberg atom arrays. Indeed, it has been demonstrated that these problems can be solved to optimality for up to thousands of nodes within minutes using both custom and generic commercial solvers on commodity hardware, without any instance-specific fine-tuning. Furthermore, this study suggests that if simulated annealing is allowed to exploit information about the underlying quasi-planar graph, its scaling performance is competitive with that of the quantum algorithm.

Some recent results have indicated that quantum annealers, when applied to large spin-glass problems that match the annealer's architecture, can be competitive against a broad range of off-the-shelf classical solvers. While these results are encouraging, further research is needed to validate their algorithmic power even in the presence of carefully tuned state-of-the-art classical algorithms. More broadly, it remains an open question whether annealers can outperform classical methods on industrially-relevant problems which do not typically have the two-dimensional structure matching the connectivity of annealers.

3 OPPORTUNITIES

Encouragingly, recent years have unveiled new and exciting opportunities for quantum algorithms in optimization. We believe that the exploration of these opportunities provides a promising path to quantum advantage.

3.1 Focus beyond polynomial-time algorithms

Polynomial-time quantum approximation algorithms have been shown to have limited performance in solving exponentially hard optimization problems. Specifically, there have been several results showing limitations and theoretical bounds on the performance of polynomial-time approximation algorithms in combinatorial optimization and first-order methods in convex optimization. Even quantum accelerations of numerical linear algebra have been proven cumbersome for speeding up polynomial-time algorithms, such as interior-point methods. Although evidence suggests potential quantum speedups for structured convex programs, such problem formulations are not prevalent in finance.

In practice, industry typically solves problems using algorithms with exponential classical complexity, such as branch-and-bound. Due to risk aversion and regulatory mandates, exact solutions are typically favored over polynomial-time approximations, even though obtaining them may require exponential runtime in the worst case. Moreover, topological obstructions based on the Overlap Gap Property apply to quantum algorithms that succeed in polynomial time with at least inverse polynomial success probability, and thus do not apply to potentially exponential-time algorithms. As a consequence, quantum speedups for exponential-time exact optimization algorithms are likely to have a significant impact on industry and provide a fruitful and underexplored research area. As an example, quantum approximate optimization algorithm (QAOA) with fixed-parameter schedules has shown promise in producing faster exponential-time algorithms [12]. Moreover, the observed advantage is significantly larger than what has been observed with prior techniques, such as short-path algorithm or amplitude amplification. A further benefit of considering exponential-time algorithms is that the scaling analysis becomes easier numerically, as the rapid growth of complexity with problem size makes even small instances hard, thus enabling empirical studies at moderate scales.

3.2 New evidence for the power of known heuristics

The mechanism and scaling performance of even well-studied quantum heuristics for optimization are not well-understood. However, recent years saw large progress, enabled by both new analytical techniques and improved simulation algorithms that are better capable to take advantage of high-performance computing systems. One promising and well-studied quantum heuristic for optimization problems is the QAOA. In QAOA, an initial quantum state is

evolved towards the solution by applying a phase operator and a mixing operator in alternation, with the parameters for the operators either fixed or optimized using a classical routine. QAOA has relatively low hardware requirements and can be benchmarked at non-trivial scale on near-term quantum hardware [14]. Recently developed analytical forms for expected solution quality of high-depth QAOA have been used to derive analytically optimal QAOA parameters and to connect optimal QAOA parameters for weighted problems to unweighted ones [16]. These results provide a broadly-applicable parameter setting heuristics, reducing the challenge of QAOA parameter optimization.

Classical simulation has been one of the main drivers of progress in understanding of QAOA. For example, recent numerical evidence shows that the insights from quantum adiabatic computing still applies in the small depth region. Specifically, QAOA performs better when the initial state is prepared to be close or exactly equal to the ground state of the mixing Hamiltonian, i.e., when the initials state aligns with the mixer [5]. Recent advances in high-performance simulation of QAOA enable scaling analysis at problem sizes that were previously unreachable. QOKit [9] implements fast simulation of high-depth QAOA by precomputing the diagonal phase operator on GPU and distributing computation across many nodes using MPI. Scaling this simulator to up to $1,024$ GPUs enables simulation of QAOA with up 40 qubits and tens of layers and provides evidence of a speedup over state-of-the-art classical methods [12]. We are optimistic that new ideas in simulation, including approximate simulation, will give new insights into performance of algorithms like QAOA at scale.

3.3 New quantum mechanisms for optimization

In addition to improving our understanding of existing heuristics, it is essential to develop new mechanisms for quantum optimization. One such approach leverages so-called "shortcuts to adiabaticity". While preparing the ground state of optimization Hamiltonian remains difficult with standard adiabatic approach due to its requirement in long evolution time, various shortcuts-to-adiabaticity techniques have been proposed to minimize losses due to diabatic transitions during the evolution within a fixed total time. In particular, counterdiabatic (CD) driving that leverages adiabatic gauge potential has been adopted to suppress diabatic transitions. By adding a velocity-dependent ancillary control field, the system can remain in the instantaneous ground state.

Although the adiabatic gauge potential theoretically exists, its practical acquisition and application in many-body systems present a considerable challenge. This is due to the requirement for detailed knowledge of the spectral properties of instantaneous Hamiltonians and the need to manipulate highly non-local multi-body interactions. To tackle this challenge, local CD driving protocols [8] have been proposed to suppress diabatic transitions with accessible local controls. These shortcuts-to-adiabaticity techniques are readily applicable in optimization problems, such as portfolio optimization, aiming for an approximate ground state with high fidelity. We remark that the combination of adiabatic evolution, CD driving, and quantum optimal control theory opens up new possibilities in solving quantum optimization challenges.

3.4 Direct integration of constraint enforcement into quantum algorithms

Production-relevant optimization problems in the industry often come with constraints. In finance, optimization problems are typically highly-constrained due to budget, risk tolerance, and regulatory requirements. Such constraints can be enforced by adding a penalty to the objective. However, the penalty strength is usually difficult to choose and the penalized objective may have undesirable properties [10], e.g. a vanishing gap in adiabatic quantum algorithms. In variational algorithms, the constraint can be incorporated into the objective for parameter optimization [4]. However, this approach still suffers from all the challenges associated with applying variational algorithms discussed above.

A promising approach explored by a number of recent results aim is to accommodate the constraints by directly modifying the quantum evolution. In QAOA, it has been shown that the evolution can be efficiently restricted to respect binary-variable constraints by applying specific mixers [5, 10]. Quantum Zeno dynamics uses repeated projective measurements to enforce that the quantum state remains in the in-constraint subspace throughout the evolution [7]. Quantum constrained Hamiltonian optimization enforces the constraint Hamiltonian and adiabatically "rotates" the objective Hamiltonian to interpolate between the worst and best feasible states while remaining in the in-constraint subspace throughout the evolution [11]. All of these techniques have been shown to lead to dramatically better performance than the penalty method. The need to incorporate more complex constraints and reduce the overhead of doing so provides an opportunity for exciting future research.

3.5 Recent advances prompt a rethink of overheads of error correction

Most of the algorithms discussed above require a fault-tolerant (FT) quantum processor to realize the quantum speedup. Fault-tolerant quantum computing involves executing fault-tolerant operators on encoded qubits, which introduces a considerable overhead. Additionally, quantum optimization algorithms heavily rely on non-Clifford gates, such as arbitrary rotation gates, whose fault-tolerant implementation of such gates is known to be challenging. To address these challenges, the community has investigated techniques such as gate synthesis, magic state distillation, and post-selection approaches. Several papers have conducted resource estimations for FT quantum optimization algorithms. Given that the research and development of FT quantum computing is still in its early stages, there exists a significant improvement space for the efficiency of fault-tolerantly executing quantum optimization algorithms, including both the provable ones and quantum heuristics.

Of particular interest are novel "good" quantum LDPC codes with constant rate (logical per physical qubits) and linear distance (growing with the number of physical qubits). Such codes have the potential to drastically lower the overhead of error correction, making at least some of the low-degree polynomial speedups practical [13]. This analysis is made urgent by the rapid progress in quantum hardware. For example, the atom shuttling in atomic systems has enabled long-range connectivity and geometrically non-local parity checks.

3.6 Quantum-inspired algorithms

While quantum algorithms for optimization still have a long way to go before they are commercially useful, quantum-inspired metaheuristic algorithms, e.g., annealing algorithms executed in Ising machines, have found some success in solving large scale binary quadratic optimization problems. However, questions regarding the propensity of these physics-based algorithms to be trapped in suboptimal solutions or their convergence rates are still outstanding. Both theoretical and empirical studies of such algorithms will help bridge the gap between the state-of-the-art classical computing hardware and quantum computers.

4 CONCLUSION

Quantum optimization stands out as one of the most promising application domains for quantum computers. While we believe quantum computing will indeed revolutionize optimization, several challenges must be overcome before we can fully realize its benefits. The vibrant quantum optimization research community is well-positioned to overcome these challenges. We now highlight three threads that appear throughout this paper, and encourage the researchers to address them in their work.

First, rigorous accumulation of evidence for scaling advantage of heuristics is of great importance. Quantum annealing community has established a template for studies of heuristics, and has some hard-earned lessons for the rest of quantum optimization community. With the increasing maturity of digital devices and the development of novel simulation techniques, benchmarking heuristics at larger scales becomes more feasible, making these lessons increasingly applicable. We remark that such evidence of speedup is valuable even if obtained in a vary narrow setting (e.g., on only a single problem) as it contributes to the accumulation of knowledge. In aggregate, many such results would remove any doubts about the potential of a given heuristic to provide a speedup. Moreover, such results will allow the community to rule out unpromising heuristics.

Second, resource estimation must be kept up-to-date with the evolving hardware landscape. The significant progress in both quantum hardware and error-correcting codes in recent years has been remarkable. For example, advancements in atomic quantum computers and error-correcting codes may shift the frontier for the kinds of speedups that are relevant in practice. Developments in software packages for resource estimation greatly simplify such work. Also, the deep integration between error correction and algorithms holds promise for significant resource reductions. The contributions from design automation community are crucial to the success of such efforts.

Finally, the development of novel mechanisms for quantum optimization and a better understanding of existing ones are necessary for achieving significant speedups. It is unlikely that integrating amplitude amplification into a classical algorithm will lead to a practical quantum speedup in the medium-term, no matter how mathematically sophisticated the construction is. Exciting recent developments make us optimistic that breakthroughs are within reach. Such opportunities include optimization algorithms based on various kinds of Hamiltonian dynamics (for both discrete and continuous optimization), the use of counterdiabatic terms, and short-path algorithms, to name just a few. The growing size and diversity of the researchers in the community are certain to lead to even more exciting ideas.

The ubiquity of optimization problems presents both challenges and opportunities. While classical solvers have been extensively optimized to tackle the problems that underlie daily operations of our economy, there remains a plethora of industrially-relevant problems that require hours to solve even with best classical solvers. Even a moderate speedup in solving these problems could lead to significant economic benefits. We hope this paper helps to focus researchers' attention on the challenges and opportunities that are most pertinent to realizing such speedups.

REFERENCES

[1] R. S. Andrist, M. J. Schuetz, P. Minssen, R. Yalovetzky, S. Chakrabarti, D. Herman, N. Kumar, G. Salton, R. Shaydulin, Y. Sun, et al. Hardness of the maximum-independent-set problem on unit-disk graphs and prospects for quantum speedups. *Physical Review Research*, 5(4):043277, 2023.

[2] S. Chakrabarti, P. Minssen, R. Yalovetzky, and M. Pistoia. Universal quantum speedup for branch-and-bound, branch-and-dut, and tree-search algorithms. *arXiv preprint arXiv:2210.03210*, 2022.

[3] E. Fontana, D. Herman, S. Chakrabarti, N. Kumar, R. Yalovetzky, J. Heredge, S. H. Sureshbabu, and M. Pistoia. The adjoint is all you need: Characterizing barren plateaus in quantum ansätze. *arXiv preprint arXiv:2309.07902*, 2023.

[4] T. Hao, R. Shaydulin, M. Pistoia, and J. Larson. Exploiting in-constraint energy in constrained variational quantum optimization. In *2022 IEEE/ACM Third International Workshop on Quantum Computing Software (QCS)*. IEEE, Nov. 2022.

[5] Z. He, R. Shaydulin, S. Chakrabarti, D. Herman, C. Li, Y. Sun, and M. Pistoia. Alignment between initial state and mixer improves QAOA performance for constrained optimization. *npj Quantum Information*, 9(1):121, 2023.

[6] D. Herman, C. Googin, X. Liu, Y. Sun, A. Galda, I. Safro, M. Pistoia, and Y. Alexeev. Quantum computing for finance. *Nature Reviews Physics*, 5(8):450–465, 2023.

[7] D. Herman, R. Shaydulin, Y. Sun, S. Chakrabarti, S. Hu, P. Minssen, A. Rattew, R. Yalovetzky, and M. Pistoia. Constrained optimization via quantum zeno dynamics. *Communications Physics*, 6(1):219, 2023.

[8] C. Li, J. Shen, R. Shaydulin, and M. Pistoia. Quantum counterdiabatic driving with local control. *arXiv preprint arXiv:2403.01854*, 2024.

[9] D. Lykov, R. Shaydulin, Y. Sun, Y. Alexeev, and M. Pistoia. Fast simulation of high-depth QAOA circuits. In *Proceedings of the SC'23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*, pages 1443–1451, 2023.

[10] P. Niroula, R. Shaydulin, R. Yalovetzky, P. Minssen, D. Herman, S. Hu, and M. Pistoia. Constrained quantum optimization for extractive summarization on a trapped-ion quantum computer. *Scientific Reports*, 12(1):1–14, 2022.

[11] M. A. Perlin, R. Shaydulin, B. P. Hall, P. Minssen, C. Li, K. Dubey, R. Rines, E. R. Anschuetz, M. Pistoia, and P. Gokhale. Q-CHOP: Quantum constrained hamiltonian optimization. *arXiv preprint arXiv:2403.05653*, 2024.

[12] R. Shaydulin, C. Li, S. Chakrabarti, M. DeCross, D. Herman, N. Kumar, J. Larson, D. Lykov, P. Minssen, Y. Sun, et al. Evidence of scaling advantage for the quantum approximate optimization algorithm on a classically intractable problem. *arXiv preprint arXiv:2308.02342*, 2023.

[13] R. Shaydulin, M. A. Perlin, C. Ryan-Anderson, S. Chakrabarti, and M. Pistoia. Focus beyond surface code for error-corrected quantum advantage. *ASCR Workshop on Basic Research Needs in Quantum Computing and Networking*, 2023.

[14] R. Shaydulin and M. Pistoia. QAOA with $N \cdot p \geq 200$. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 01, pages 1074–1077, 2023.

[15] K. Shi, R. Herrman, R. Shaydulin, S. Chakrabarti, M. Pistoia, and J. Larson. Multi-angle QAOA does not always need all its angles. In *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*. IEEE, Dec. 2022.

[16] S. H. Sureshbabu, D. Herman, R. Shaydulin, J. Basso, S. Chakrabarti, Y. Sun, and M. Pistoia. Parameter setting in quantum approximate optimization of weighted problems. *Quantum*, 8:1231, 2024.