QUBO formulation of mobile communication scheduling for quantum computers

Combinatorial Optimization with Hybrid Quantum-Classical Algorithms

Olli Apilo • Erik Hieta-aho • Jorma Kilpi • Hannu Reittu • Kari Seppänen

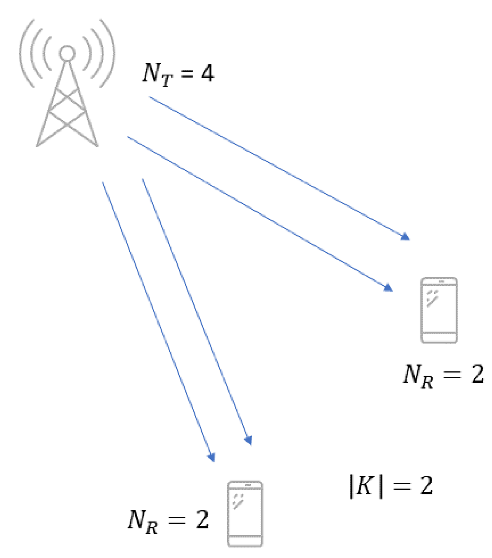
Introduction

Noisy Intermediate Scale Quantum (NISQ) devices have empirical evidence of quantum computational supremacy as scale-up of a search space rather than speed-up performance [3]. In the context of wireless communications, we study combinatorial optimization problems that have a huge search space.

One problem with a huge search space is the **Multi-User Multiple Input Multiple Output (MU-MIMO)** downlink scheduling problem of 5G and future 6G networks. There exist classical algorithms [2] that are very fast in solving the MU-MIMO scheduling problem, but with strong heuristics and parallel computing. A quantum algorithm for the problem may allow to study the full search space and provide insight for even better classical algorithms.

The MU-MIMO downlink scheduling problem

A base station (BS) must share its limited bandwidth fairly between a large number of **user equipments (UEs)**. Scheduler selects which UEs are multiplexed on a given time-frequency resource block (RB), how many data streams per UE are transmitted, and which modulation and coding scheme (MCS) is used for these UEs.



MU-MIMO downlink scheduling problem.

Case study: SU-MIMO DL scheduling

It is not straightforward to formulate the general MU-MIMO DL scheduling problem as QUBO because of the signal-to-interference-plus-noise ratio (SINR) dependency between the scheduled UEs. As a first approach, we have focused on the simpler single-user (SU) MIMO problem where an RB is scheduled to only one UE. The target is to maximize the sum rate subject to realistic constraints in 5G systems.

Definitions

- B: set of RBs to be allocated, $|B| \leq 273$
- K: set of UEs to be scheduled, typically $|K| < 100$
- M: set of possible MCSs, $|M| = 29$
- N_T: max. number of data streams BS can transmit, $N_T \leq 8$
- N_R: max. number of data streams a UE can receive, $N_R \leq 4$, $N_R \leq N_T$
- r_m: predefined number data bits per RB given MCS m
- $\gamma_{k,b,y}$: SINR at RB b for UE k with y data streams
- θ^m: SINR threshold at which a data block can be successfully decoded with MCS m
- $r_{k,b,y}$: number data bits at RB b for UE k with y data streams
- y_k: number of scheduled data streams for user k
- $x_{k,b}$: binary decision variable indicating if RB b is scheduled for user k
- $w_{k,y}$: binary decision variable indicating if the number of data streams is y for user k
- $z_{k,m}$: binary decision variable indicating if MCS m is assigned for user k

Optimization problem

If the SINR is known and MCS m is selected, $r_{k,b,y}$ is given as

$$r_{k,b,y} = \begin{cases} r_m, & \gamma_{k,b,n} \geq \theta_m, \\ 0, & \text{otherwise} \end{cases}$$

Maximize the sum rate $y_k r_{k,b,y} x_{k,b} w_{k,y} z_{k,m}$ subject to

1. $\sum_{k \in K} x_{k,b} \leq 1,\ \forall b \in B$: only one UE can be scheduled per RB
2. $\sum_{i=1}^{N_R} w_{k,i} \leq 1,\ \forall k \in K$: only one number of data streams can be used for a UE over its scheduled RBs
3. $\sum_{m \in M} z_{k,m} \leq 1,\ \forall k \in K$: only one MCS can be used for a UE over its scheduled RBs

QUBO formulation

Let vector **a** include all the binary decision variables $x_{k,b}$, $w_{k,y}$, and $z_{k,m}$. In the worst case, length of a is $|K|(|B|+N_R+|M|)$ The number of decision variables can be reduced by mapping only relevant $z_{k,m}$:

- Smallest relevant m fulfills $\theta_m \leq \min\{SINR_k^1, \dots, SINR_k^B\} < \theta_{m+1}$
- Largest relevant m fulfills $\theta_m \leq \max\{SINR_k^1, \dots, SINR_k^B\} < \theta_{m+1}$
- Relevant MCS subset per user: $M_k \subseteq M$

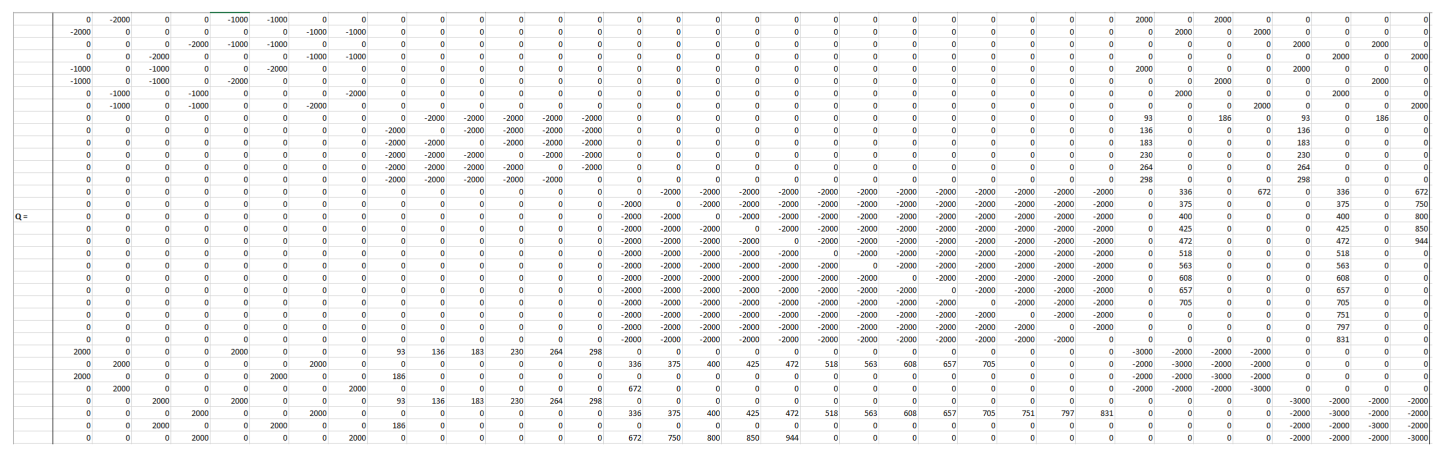
The third order sum-rate optimization problem is converted to QUBO using Rosenberg's procedure [1] with the cost of introducing $|B||K|N_R$ additional variables. An example mapping of variables when $|B| = |K| = N_R = 2$:

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}
$x_{1,1}$	$x_{2,1}$	$x_{1,2}$	$x_{2,2}$	$w_{1,1}$	$w_{1,2}$	$w_{2,1}$	$w_{2,2}$	$z_{1,8}$	$z_{1,9}$	$z_{2,9}$	$z_{2,10}$	$x_{1,1}w_{1,1}$	$x_{2,1}w_{2,1}$	$x_{1,1}w_{1,2}$	$x_{2,1}w_{2,2}$	$x_{1,2}w_{1,1}$	$x_{2,2}w_{2,1}$	$x_{1,2}w_{1,2}$	$x_{2,2}w_{2,2}$

Constraints are transformed to quadratic penalties, e.g. for Constraint 1

$$\sum_{i=1}^{|K||B|} a_i \leq 1 \quad \Rightarrow \quad \sum_{i=1}^{|K||B|-1} \sum_{j=i+1}^{|K||B|} E(a_i a_j + a_j a_i)$$

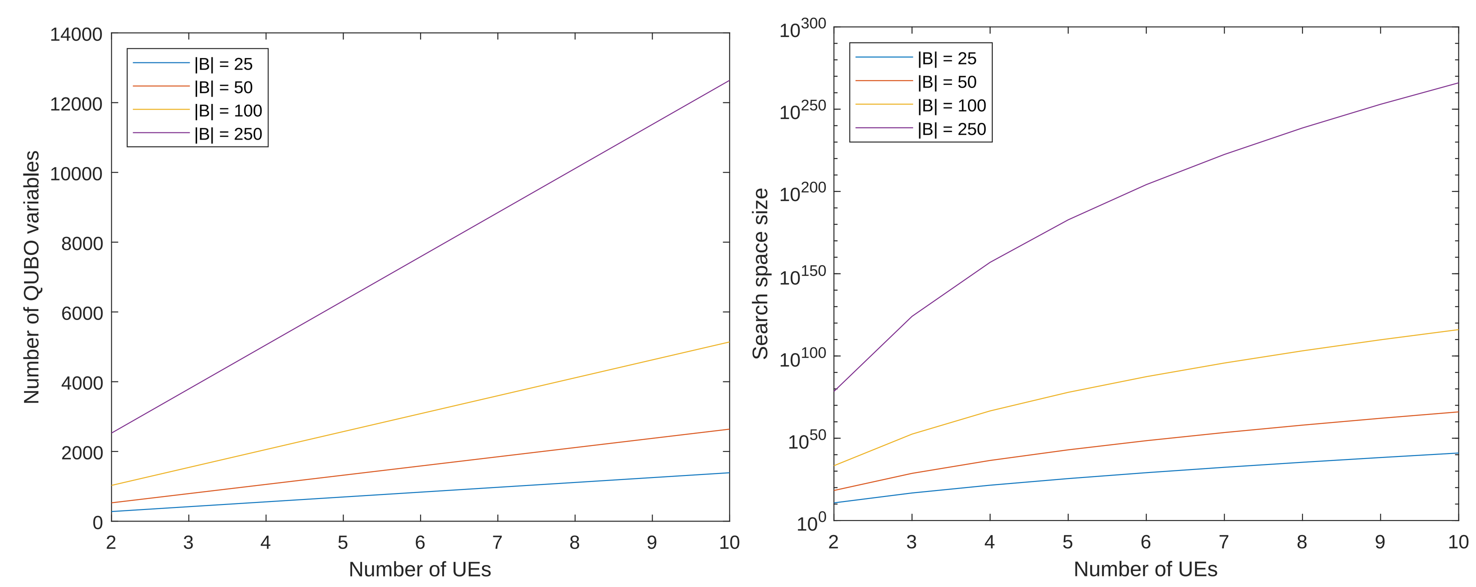
QUBO matrix is filled with $r_{k,b,y}$ values, penalty terms from Rosenberg's quadratization, and penalty terms from constraints.



Example QUBO matrix with $|B| = |K| = N_R = 2$, $|M_1| = 6$, $|M_2| = 13$.

The number of QUBO variables: $|B||K| + |K|N_R + \sum_{i=1}^{|K|} |M_k| + |B||K|N_R$

Worst case search space size: $|K|^{|B|} N_R^{|K|} \left(\prod_{i=1}^{|K|} |M_k| \right)$, $|B| > |K|$



Number of QUBO variables and search space size when $N_R = 4$ and $|M_k| = 10,\ \forall k \in K$.

Next steps

Our plan is to study the MU-MIMO scheduling problem in several different ways including the following:

- Consider other scheduling criteria than sum rate.
- Attempt to find alternative ways for QUBO formulation (with less variables).
- Extend study to MU-MIMO scheduling.
- Study the possibilities to implement the QUBO problem on a quantum computer with quantum annealing or gate-based QAOA.

Our study is funded by Business Finland, see `https://www.cohqca.fi/` for further information. Companies in the project steering group are Nokia Bell Labs, Unitary Zero Space and Cumucore.

Unitary Zero Space

Cumucore

VTT

Summary

- **We are focusing on MU-MIMO downlink scheduling problem, which has small input, small output, and an enormous search space.**
- **As a first approach, we have formulated a QUBO problem for maximizing the sum rate of SU-MIMO downlink and verified its correctness against reference brute force solutions for cases with reasonably small search space.**
- **With the QUBO formulation, we can proceed with purely quantum, or quantum inspired classical algorithms.**

References

[1] Martin Anthony, Endre Boros, Yves Crama, and Aritanan Gruber. Quadratic reformulations of nonlinear binary optimization problems. *Mathematical Programming*, 162:115–144, June 2017.

[2] Yongce Chen, Yubo Wu, Y. Thomas Hou, and Wenjing Lou. mCore+: A real-time design achieving ∼ 500 μs scheduling for 5G MU-MIMO systems. *IEEE Transactions on Mobile Computing*, 22(12):7249–7265, December 2023.

[3] Youngseok Kim et al. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023.

Contact: Kari Seppänen, Senior Scientist
Jorma Kilpi, Senior Scientist
Firstname.Lastname@vtt.fi