arXiv:2307.11160v2 [quant-ph] 5 Mar 2024

# Time-optimal multi-qubit gates: Complexity, efficient heuristic and gate-time bounds

P. Baßler[1], M. Heinrich[1], and M. Kliesch[2]

[1] Institute for Theoretical Physics, Heinrich Heine University Düsseldorf, Germany

[2] Institute for Quantum Inspired and Quantum Optimization, Hamburg University of Technology, Germany

Multi-qubit entangling interactions arise naturally in several quantum computing platforms and promise advantages over traditional two-qubit gates. In particular, a fixed multi-qubit Ising-type interaction together with single-qubit X-gates can be used to synthesize global ZZ-gates (GZZ gates). In this work, we first show that the synthesis of such quantum gates that are time-optimal is NP-hard. Second, we provide explicit constructions of special time-optimal multi-qubit gates. They have constant gate times and can be implemented with linearly many X-gate layers. Third, we develop a heuristic algorithm with polynomial runtime for synthesizing fast multi-qubit gates. Fourth, we derive lower and upper bounds on the optimal GZZ gate-time. Based on explicit constructions of GZZ gates and numerical studies, we conjecture that any GZZ gate can be executed in a time $\mathrm{O}(n)$ for $n$ qubits. Our heuristic synthesis algorithm leads to GZZ gate-times with a similar scaling, which is optimal in this sense. We expect that our efficient synthesis of fast multi-qubit gates allows for faster and, hence, also more error-robust execution of quantum algorithms.

## 1 Introduction

Any quantum computation requires to decompose its logical operations into the platform's native instruction set. The performance of the computation depends heavily on the available instructions and their implementation in the quantum hardware. In particular for early quantum devices a major challenge is posed by their short decoherence times, which limits the runtime of a quantum computation significantly. Therefore, it is not only necessary to optimize the number of native instructions but also their execution time.

Ising-type interactions give rise to an important and rich class of Hamiltonians ubiquitous in several quantum computing platforms [1–5]. Previously, we have utilized these Ising-type interactions in a new synthesis method [6]. In particular, we have considered the problem of synthesizing time-optimal multi-qubit gates on a quantum computing platform that supports the following basic operations:

(I) single-qubit rotations can be executed in parallel, and

(II) it offers a fixed Ising-type interaction.

The corresponding synthesis of global ZZ-gates (GZZ gates) is given by the minimization of the overall gate time, which can be written as a linear program (LP) [6]. This LP is exponentially large in the number of qubits. It has been unclear whether such multi-qubit gates can be synthesized in a computationally efficient way while keeping the gate time optimal.

In this work, we prove that this synthesis problem is NP-hard and provide a close-to-optimal efficient heuristic solution. To establish this hardness result, we draw the connection between the synthesis problem and graph theory. The *cut polytope* is defined as the convex hull of binary vectors representing the possible cuts of a given graph [7]. We provide a polynomial time reduction of the

P. Baßler: bassler@hhu.de
M. Kliesch: martin.kliesch@tuhh.de

Accepted in Quantum 2024-03-04, click title to verify.
Published under CC-BY 4.0.

membership problem of the cut polytope to the synthesis of time-optimal multi-qubit gates. Since this membership problem is NP-complete [8], the synthesis of time-optimal multi-qubit gates is NP-hard. This is akin to the NP-hardness of finding an optimal control pulse for multi-qubit gates using the Mølmer-Sørensen mechanism [4, 5].

We provide several ways to circumvent the hardness of time-optimal GZZ-gates. First, we provide explicit constructions of time-optimal multi-qubit gates realizing nearest-neighbor coupling under physically motivated assumptions. Such constructed nearest-neighbor multi-qubit gates exhibit a constant gate time and can be implemented with only linearly many single-qubit gate layers. We then use these ideas to define relaxations of the underlying linear program, leading to a hierarchy of polynomial-time algorithms for the synthesis of fast multi-qubit gates. By increasing the level in the hierarchy, this heuristic approach can be adapted to provide substantially better approximations to the optimal solution at the cost of higher polynomial runtime. For a small number of qubits, numerical experiments show that the so-obtained gate times are close to the optimal solution and come with significant runtime savings.

Among others, we prove bounds on the minimal multi-qubit gate time, and conjecture that it scales at most linear with the number of qubits. This claim is supported by a class of explicit constructions of time-optimal multi-qubit gates achieving the linear upper bound. Moreover, we provide numerical evidence that these explicit solutions in fact yield the longest gate time for a small number of qubits.

We expect our results to be useful for the implementation of time-optimal multi-qubit gates in noisy and intermediate scale quantum (NISQ) devices and beyond. The polynomial-time heuristic algorithm makes it possible to efficiently synthesize fast multi-qubit gates for a growing number of qubits. The here considered multi-qubit gates have been useful in a number of different applications, some of which we investigated in previous work [6]. Furthermore, they are the main building blocks for a class of Instantaneous Quantum Polynomial time (IQP) circuits which might be classically hard to simulate [9]. More recently, it was shown that quantum memory circuits and boolean functions can be implemented with a constant number of GZZ gates and additional ancilla qubits [10]. Moreover, there is an ancilla-free construction of multi-qubit Clifford circuits using at most 26 GZZ gates [11]. As noted in Ref. [11], there is also a shorter implementation for $n \leq 2^{12}$, requiring only $2(\log_2(n) + 1)$ GZZ gates. This implementation is based on a decomposition in Ref. [12] and the log-depth implementation of a CX circuit in formula (4) of Ref. [13].

This paper is structured as follows: We first give a brief introduction to the synthesis of time-optimal multi-qubit gates [6]. In Section 4 we proof that the time-optimal multi-qubit synthesis problem is NP-hard. However, in Section 5 we explicitly construct a certain class of time-optimal multi-qubit gates with constant gate time. The heuristic algorithm based on the ideas of the previous section is introduced in Section 6. Section 7 provides gate time bounds for time-optimal multi-qubit gates. Finally, Section 8 presents numerical evidence for the conjectured linear gate-time scaling and numerical benchmarks for the heuristic algorithm.

## 2 Synthesizing multi-qubit gates with Ising-type interactions

In this section, we give a short introduction to the synthesis of time-optimal multi-qubit gates. For more details we refer to the first two sections of Ref. [6].

On the abstract quantum computing platform with $n$ qubits specified by the requirements (I) and (II) above, interactions between the qubits are generated by an Ising-type Hamiltonian

$$H_{ZZ} \coloneqq -\sum_{i<j}^{n} J_{ij}\sigma_z^i\sigma_z^j\,, \tag{1}$$

where $\sigma_z^i$ is the Pauli-Z operator acting on the $i$-th qubit. Note, that diagonal terms, where $i = j$, are excluded since they only change the Hamiltonian by an energy offset. By $J$ we denote the symmetric matrix with entries $J_{ij}$ in the upper triangular part and vanishing diagonal. We call $J$ the *physical coupling matrix*.

Conjugating the Hamiltonian $H_{ZZ}$ with X gates on the qubits indicated by the binary vector

$\boldsymbol{b} \in \mathbb{F}_2^n$ yields

$$\begin{aligned}
\sigma_x^{\boldsymbol{b}} H_{ZZ} \sigma_x^{\boldsymbol{b}} &= -\sum_{i<j}^{n} J_{ij} \sigma_x^{b_i} \sigma_x^{b_j} \sigma_z^i \sigma_z^j \sigma_x^{b_i} \sigma_x^{b_j} \\
&= -\sum_{i<j}^{n} J_{ij} (-1)^{b_i} (-1)^{b_j} \sigma_z^i \sigma_z^j \\
&= -\sum_{i<j}^{n} J_{ij} m_i m_j \sigma_z^i \sigma_z^j =: H(\boldsymbol{m}) \,,
\end{aligned} \tag{2}$$

where we define the *encoding* $\boldsymbol{m} \coloneqq (-1)^{\boldsymbol{b}}$ entry wise. The sign of the interaction between qubit $i$ and $j$ is given by $m_i m_j \in \{-1, +1\}$. We call $H(\boldsymbol{m})$ the *encoded Hamiltonian*.

Given time steps $\lambda_{\boldsymbol{m}} \geq 0$ during which the encoding $\boldsymbol{m}$ is used, we consider unitaries of the form

$$\prod_{\boldsymbol{m}} \mathrm{e}^{-\mathrm{i}\lambda_{\boldsymbol{m}} H(\boldsymbol{m})} = \mathrm{e}^{-\mathrm{i}\sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} H(\boldsymbol{m})} =: \mathrm{e}^{-\mathrm{i}H} \,, \tag{3}$$

where we used that the diagonal Hamiltonians $H(\boldsymbol{m})$ mutually commute. For all possible encodings $\boldsymbol{m} \in \{-1, +1\}^n$ we collect the time steps $\lambda_{\boldsymbol{m}}$ in a vector $\boldsymbol{\lambda} \in \mathbb{R}^{2^n}$ and interpret $t = \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}}$ as the total gate time of the unitary $\mathrm{e}^{-\mathrm{i}H}$, implemented by the sequence of unitaries (3). Moreover, we use the symmetry

$$(-\boldsymbol{m})(-\boldsymbol{m})^T = \boldsymbol{m}\boldsymbol{m}^T \tag{4}$$

to reduce the degrees of freedom in $\boldsymbol{\lambda}$ from $2^n$ to $2^{n-1}$ by adding up $\lambda_{\boldsymbol{m}} + \lambda_{-\boldsymbol{m}}$ to a single time step.

The so generated unitary is the time evolution operator under the *total Hamiltonian*

$$H = -\sum_{i<j}^{n} A_{ij} \sigma_z^i \sigma_z^j \,, \tag{5}$$

where we have defined the *total coupling matrix*

$$A \coloneqq J \circ \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} \boldsymbol{m}\boldsymbol{m}^T \,, \tag{6}$$

and used the linearity of the Hadamard (entry-wise) product $\circ$. By construction, $A$ is a symmetric matrix with vanishing diagonal. Let us define the $\binom{n}{2}$-dimensional subspace of symmetric matrices with vanishing diagonal by

$$\mathrm{Sym}_0(\mathbb{R}^n) \coloneqq \{M \in \mathrm{Sym}(\mathbb{R}^n) \mid M_{ii} = 0 \ \forall i \in [n]\} \,. \tag{7}$$

For $A \in \mathrm{Sym}_0(\mathbb{R}^n)$, we define an associated multi-qubit gate $\mathsf{GZZ}(A)$, where $\mathsf{GZZ}$ stands for "global $ZZ$ interactions",

$$\mathsf{GZZ}(A) \coloneqq \exp\left(\mathrm{i} \sum_{i<j}^{n} A_{ij} \sigma_z^i \sigma_z^j\right) \,. \tag{8}$$

To determine which matrices can be decomposed as in Eq. (6), we denote the non-zero index set of a symmetric matrix as $\mathrm{nz}(M) \coloneqq \{(i,j) \mid M_{ij} \neq 0, i < j\}$. Then, the subspace of matrices $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ that can be decomposed as in Eq. (6) is exactly given by the condition $\mathrm{nz}(A) \subseteq \mathrm{nz}(J)$, which we assume to hold from now on. Thus, all-to-all connectivity enables to decompose any coupling matrix $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ but is not strictly required by our approach. We call the number of encodings $\boldsymbol{m}$ needed for the decomposition the *encoding cost* of $\mathsf{GZZ}(A)$, and $\sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}}$ the *total* $\mathsf{GZZ}$ *time*. Note, that both quantities depend on the chosen decomposition.

It is convenient to abstract the following analysis from the physical details given by $J$. For a matrix $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ with $\mathrm{nz}(A) \subseteq \mathrm{nz}(J)$, its possible decompositions are in one-to-one correspondence with the decompositions of the matrix $M \coloneqq A \oslash J \in \mathrm{Sym}_0(\mathbb{R}^n)$ where

$$(A \oslash J)_{ij} \coloneqq \begin{cases} A_{ij}/J_{ij}, & i \neq j \text{ and } J_{ij} \neq 0 \,, \\ 0, & \text{otherwise} \,. \end{cases} \tag{9}$$

We further define the linear operator

$$\begin{aligned}\mathcal{V} : \mathbb{R}_{\geq 0}^{2^{n-1}} &\to \mathrm{Sym}_0(\mathbb{R}^n), \\ \boldsymbol{\lambda} &\mapsto \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} \boldsymbol{m}\boldsymbol{m}^T,\end{aligned} \tag{10}$$

represented in the standard basis by a matrix

$$V \in \{-1, +1\}^{\binom{n}{2} \times (2^{n-1})}. \tag{11}$$

Let $\boldsymbol{v} : \mathrm{Sym}_0(\mathbb{R}^n) \to \mathbb{R}^{\binom{n}{2}}$ be the (row-wise) vectorization of the upper triangular part of the matrix input such that the columns of $V$ are given by $\boldsymbol{v}(\boldsymbol{m}\boldsymbol{m}^T)$. Our objective is to minimize the total gate time and the amount of $\boldsymbol{m}$'s needed to express the matrix $M \in \mathrm{Sym}_0(\mathbb{R}^n)$. To this end we formulate the following linear program (LP):

$$\begin{aligned}\text{minimize} \quad & \mathbf{1}^T \boldsymbol{\lambda} \\ \text{subject to} \quad & V\boldsymbol{\lambda} = \boldsymbol{v}(M), \\ & \boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^{2^{n-1}},\end{aligned} \tag{12}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$ is the all-ones vector such that $\mathbf{1}^T\boldsymbol{\lambda} = \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}}$. A *feasible solution* is an assignment of the variables that fulfills all constraints of the optimization problem and an *optimal solution* is a feasible solution which also minimizes/maximizes the objective function. Throughout this paper, we indicate optimal solutions by an asterisk, e.g. $\boldsymbol{\lambda}^*$. Note, that the LP (12) has a feasible solution for any symmetric matrix $M$ with vanishing diagonal [6, Theorem II.2]. The theory of linear programming then guarantees the existence of an optimal solution with at most $\binom{n}{2}$ non-zero entries [6, Proposition II.3].

A standard tool in convex optimization is *duality* [14] which will be used in Section 7. The dual LP to the LP (12) reads as follows:

$$\begin{aligned}\text{maximize} \quad & \langle M, \boldsymbol{y} \rangle \\ \text{subject to} \quad & V^T\boldsymbol{y} \leq \mathbf{1}, \\ & \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}},\end{aligned} \tag{13}$$

with the inner product $\langle \cdot, \cdot \rangle : \mathrm{Sym}_0(\mathbb{R}^n) \times \mathbb{R}^{\binom{n}{2}} \to \mathbb{R}$, $\langle M, \boldsymbol{y} \rangle \mapsto \boldsymbol{v}(M)^T\boldsymbol{y}$. Here, inequalities between vectors are to be understood entry-wise. A simple, but important fact is the following: If $\boldsymbol{\lambda}^*$ is an optimal solution to the primal LP (12), then any feasible solution $\boldsymbol{y}$ to the dual LP (13) provides a *lower bound* as $\langle M, \boldsymbol{y} \rangle \leq \mathbf{1}^T\boldsymbol{\lambda}^*$. Moreover, the feasibility of the LP (12) implies that we have *strong duality*: if $\boldsymbol{y}^*$ is a dual optimal solution, then we have equality between the optimal values, $\langle M, \boldsymbol{y}^* \rangle = \mathbf{1}^T\boldsymbol{\lambda}^*$.

## 3 Main results

We want to highlight our main contributions. First, we provide the hardness results for the synthesis of time-optimal GZZ gates.

**Theorem** (Theorem 6)**.** *The decision version of the LP* (12)*, is* NP*-complete.*

We say that the synthesis of time-optimal multi-qubit gates (solving LP (12)) is NP-hard in the sense of the function problem extension of the decision problem class NP [15]. We circumvent the hardness of the time-optimal multi-qubit gate synthesis by providing an explicit construction of GZZ gates realizing next neighbor coupling with constant gate time and linear encoding cost. The assumption of a constant subdiagonal of $J$ is physically motivated and can be realized in an ion trap [16].

**Theorem** (Theorem 10, informal)**.** *Let the subdiagonal of $J$ and $A$ be a constant with values $c$ and $\varphi$ respectively. Then* $\mathsf{GZZ}(A)$ *on $n$ qubits has the encoding cost of $d \leq 2n$ and constant total* $\mathsf{GZZ}$ *time $2\varphi/c$. This total gate time is optimal.*

In Section 6 we define Algorithm 3 and introduce LP (41), a polynomial time heuristic to synthesize $\mathsf{GZZ}$ gates with small, but not necessarily minimal, gate time. This heuristic does not rely on any further assumptions and is applicable for arbitrary $J, A \in \mathrm{Sym}_0(\mathbb{R}^n)$. In Section 8 we show numerically that this heuristic leads to $\mathsf{GZZ}$ gate times close to the optimum while solving the synthesis problem much faster.

Furthermore, we proof lower and upper bounds on the $\mathsf{GZZ}$ gate time.

**Theorem** (Theorem 15)**.** *The optimal total gate time of* $\mathsf{GZZ}(A)$ *with $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ is lower and upper bounded by*

$$\|A \oslash J\|_{\ell_\infty} \leq \mathbf{1}^T \boldsymbol{\lambda}^* \leq \|A \oslash J\|_{\ell_1} . \tag{14}$$

Here, the upper bound scales quadratic in $n$ for a constant matrix $M = A \oslash J$. This upper bound is loose in the sense that it also holds for the $\mathsf{GZZ}$ gates constructed by the heuristic in Section 6. Therefore, we tighten the upper bound to a linear scaling in $n$.

**Conjecture** (Conjecture 16)**.** *The optimal gate time of* $\mathsf{GZZ}(A)$ *with $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ is tightly upper bounded by*

$$\mathbf{1}^T \boldsymbol{\lambda}^* \leq \|A \oslash J\|_{\ell_\infty} \cdot \begin{cases} n\,, & \text{for odd } n\,, \\ n-1\,, & \text{for even } n\,. \end{cases} \tag{15}$$

We provide evidence for this conjecture using an explicit construction (Theorem 21) that realizes the conjectured upper bound for any $n$, as well as numerical evidence for $n \leq 8$ (Fig. 1). Unfortunately, we were not able to prove this result and state the challenges in Appendix A.

## 4 Synthesizing time-optimal GZZ gates is NP-hard

In this section, we investigate the complexity of solving the gate synthesis problem stated as LP (12). We observe that LP (12) is an optimization over the convex cone generated by

$$\mathcal{E}_n := \left\{ \boldsymbol{m}\boldsymbol{m}^T \;\middle|\; \boldsymbol{m} \in \{-1,+1\}^n, m_n = +1 \right\}, \tag{16}$$

which is the set of outer products generated by all possible encodings $\boldsymbol{m}$. Due to the symmetry Eq. (4) we can uniformly fix the value of one entry of $\boldsymbol{m}$. We chose the convention $m_n = +1$. In the literature $\mathcal{E}_n$ is also known as the *elliptope* of rank one matrices [17]. In the following, we consider the polytope

$$\mathrm{conv}(\mathcal{E}_n) := \left\{ \sum_i \lambda_i \boldsymbol{r}_i \;\middle|\; \sum_i \lambda_i = 1, \lambda_i \geq 0, \boldsymbol{r}_i \in \mathcal{E}_n \right\}, \tag{17}$$

and show the connection to graph theory, in particular to the cuts of graphs.

**Definition 1** (cut polytope [7])**.** *Let $K_n = (V_n, E_n)$ be a complete graph with $n$ vertices. Denote $\delta(X)$ the set of all edges with one endpoint in $X \subset V_n$ and the other endpoint in its complement $\bar{X}$, i.e., $\delta(X)$ defines the cut between $X$ and $\bar{X}$. Let $\chi^{\delta(X)} \in \{0,1\}^{|E_n|}$ denote the characteristic vector of a cut, with $\chi_e^{\delta(X)} = 1$ if $e \in \delta(X)$ and $\chi_e^{\delta(X)} = 0$ otherwise. We define the* cut polytope *as the convex hull of the characteristic vectors*

$$\mathrm{CUT}_n := \mathrm{conv}\left\{ \chi^{\delta(X)} \in \{0,1\}^{|E_n|} \;\middle|\; X \subseteq V_n \right\}. \tag{18}$$

**Lemma 2.** *For all $n$, $\mathrm{CUT}_n$ is isomorphic to $\mathrm{conv}(\mathcal{E}_n)$.*

*Proof.* For each $X \subset V_n$ we set $m_i = +1$ if $i \in X$ and $m_i = -1$ if $i \in \bar{X}$. Note, that there are $2^{n-1}$ different pairs of $X$ and $\bar{X}$. We then have $m_i m_j = -1$ if $i \in X$ and $j \in \bar{X}$ (or the other way around), and $m_i m_j = +1$ if $i, j \in X$ or $i, j \in \bar{X}$. So the characteristic vector can be written as $\chi_e^{\delta(X)} = (1 - m_i m_j)/2$ for each edge $e \in E_n$ connecting vertices $i$ and $j$. This is clearly a bijective affine map between the vertices and thus $\mathrm{CUT}_n$ is isomorphic to $\mathrm{conv}(\mathcal{E}_n)$. □

The following decision problems are membership problems, where the task is to decide if a given element $\boldsymbol{x}$ belongs to a set or not. In our case $\boldsymbol{x}$ is a vector and the set is a polytope.

**Problem 3** (CUT$_n$ membership)**.**

Instance *The adjacency matrix* $M \in \mathrm{Sym}_0(\mathbb{Q}^n_{\geq 0})$ *of a weighted undirected graph with non-negative weights.*

Question *Is* $M \in \mathrm{CUT}_n$*?*

**Problem 4** (conv($\mathcal{E}_n$) membership)**.**

Instance *The matrix* $M \in \mathrm{Sym}_0(\mathbb{Q}^n)$*.*

Question *Is* $M \in \mathrm{conv}(\mathcal{E}_n)$*? That is, does there exist a decomposition* $M = \sum_i \lambda_i \boldsymbol{r}_i$ *with* $\boldsymbol{r}_i \in \mathcal{E}_n$ *such that* $\lambda_i \geq 0$ *and* $\sum_i \lambda_i = 1$*?*

It is well known that the membership problem of the cut polytope, Problem 3, and Problem 4 are NP-complete [8, 18, 19]. Next, we state the decision version of our gate synthesis optimization.

**Problem 5** (time-optimal multi-qubit gate synthesis)**.**

Instance *The matrix* $M = A \oslash J \in \mathrm{Sym}_0(\mathbb{Q}^n)$ *and a constant* $K \in \mathbb{Q}_{\geq 0}$*.*

Question *Is there a decomposition* $M = \sum_i \lambda_i \boldsymbol{r}_i$ *with* $\boldsymbol{r}_i \in \mathcal{E}_n$ *such that* $\lambda_i \geq 0$ *and* $\sum_i \lambda_i \leq K$*?*

**Theorem 6.** *Problem 5, which is the decision version of the LP (12), is* NP*-complete.*

*Proof.* A solution of Problem 5 can be verified in polynomial time since there always exists a decomposition $\sum_i \lambda_i \boldsymbol{r}_i$ of $M$ with minimal $\sum_i \lambda_i$ which has at most $\binom{n}{2} = n/2(n-1)$ non-zero terms [6, Proposition II.3]. Therefore, Problem 5 is NP.

To show that Problem 5 is NP-hard, we construct a polynomial-time mapping reduction from Problem 4 to Problem 5. Given the matrix $M \in \mathrm{Sym}_0(\mathbb{Q}^n)$ and a constant $K \in \mathbb{Q}_{\geq 0}$ as an instance for Problem 5, let $\lambda_i \geq 0$ be the positive coefficients of the decomposition. If we find $\sum_i \lambda_i < K$, then we can always add additional $\lambda$'s such that equality holds. We choose the additional $\lambda$'s as the coefficients of the decomposition for the all-zero matrix, see e.g. Lemma 8 below with $\boldsymbol{k} = \mathbf{1}_n$ for an explicit construction. We define the matrix $M' := M/K$ and the positive coefficients $\lambda'_i := \lambda_i/K$. Then $M' = \sum_i \lambda'_i \boldsymbol{r}_i$ with $\boldsymbol{r}_i \in \mathcal{E}_n$ and $\sum_i \lambda'_i = 1$. □

# 5 Time-optimal GZZ synthesis for special instances

Although, solving the LP (12) is NP-hard we present explicit optimal solutions for certain families of instances which is equivalent to constructing time-optimal GZZ gates. The constructions of this section yield a qubit-independent total GZZ time which satisfies the optimal lower bound of Lemma 7 below. Moreover, we show that some GZZ gates can be synthesized with an encoding cost independent of the number of qubits. However, most of these constructions assume constant values for the elements of the band-diagonal of the physical coupling matrix $J$. These assumptions are relaxed throughout this section providing explicit optimal solutions for physically relevant cases. These results build the foundation for the heuristic algorithm for fast GZZ gate synthesis in the next section.

By $\|M\|_{\ell_p}$ we denote the $\ell_p$-norm of a symmetric matrix $M$, which is given as the $\ell_p$-norm of a vector $\boldsymbol{v}(M)$ containing all lower/upper triangular matrix elements in some order. First, we proof a lower bound on the optimal total GZZ time which can be used to verify time-optimality.

**Lemma 7.** *For any* $M \in \mathrm{Sym}_0(\mathbb{R}^n)$ *the optimal objective function value of the LP (12) is lower bounded by*

$$\|M\|_{\ell_\infty} \leq \mathbf{1}^T \boldsymbol{\lambda}^* \,. \tag{19}$$

*Proof.* The lower bound can be verified by the fact that the matrix representation $V$ of the linear operator in Eq. (10) only has entries $\pm 1$ and that $\boldsymbol{\lambda}^*$ is non-negative. Thus, it holds that $\|M\|_{\ell_\infty} = \|V\boldsymbol{\lambda}^*\|_{\ell_\infty} \leq \mathbf{1}^T \boldsymbol{\lambda}^*$. □

Next, we provide calculation rules for the coupling matrix $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ of the $\mathsf{GZZ}(A)$ gate. These rules are inherited from matrix exponentials. Let $A_1, A_2, A_3 \in \mathrm{Sym}_0(\mathbb{R}^n)$ then

(i) $\mathsf{GZZ}(A_1 + A_2) = \mathsf{GZZ}(A_1)\,\mathsf{GZZ}(A_2)$,

(ii) $\mathsf{GZZ}(A_1 \oplus A_2) = \mathsf{GZZ}(A_1) \otimes \mathsf{GZZ}(A_2)$ and

(iii) $\mathsf{GZZ}(A_1 \circ A_2) = \mathsf{GZZ}(A_3)$, where the coupling matrices can be decomposed as

$$A_1 := \sum_{k=1}^{d_1} \lambda_{\boldsymbol{m}_k} \boldsymbol{m}_k \boldsymbol{m}_k^T, \quad A_2 := \sum_{k=1}^{d_2} \beta_{\boldsymbol{v}_k} \boldsymbol{v}_k \boldsymbol{v}_k^T, \quad A_3 := \sum_{k=1}^{d_3} \gamma_{\boldsymbol{w}_k} \boldsymbol{w}_k \boldsymbol{w}_k^T,$$

with $d_3 = d_1 d_2$, $\boldsymbol{w}_k = (\boldsymbol{m}_i \circ \boldsymbol{v}_j)_{k=d_2 i+j}$ and $\gamma_{\boldsymbol{w}_k} = (\lambda_{\boldsymbol{m}_i} \beta_{\boldsymbol{v}_j})_{k=d_2 i+j}$ for $j = 1, \ldots, d_2$ and $i = 0, \ldots, d_1 - 1$.

In Item (iii) we can also express $\boldsymbol{\gamma} = \boldsymbol{\lambda} \otimes \boldsymbol{\beta}$ with the Kronecker product. Note, that these generated $\mathsf{GZZ}$ gates are not necessarily time-optimal.

We denote the $k \times k$ identity matrix by $\mathbb{1}_k$, the $k$ dimensional all-ones vector by $\mathbf{1}_k$ and the $k \times k$ matrix of ones with vanishing diagonal by $E_k := \mathbf{1}_k \mathbf{1}_k^T - \mathbb{1}_k$. With the next two lemmas, we bound the encoding cost and total gate time for special cases of Item (i) and Item (ii), where the total gate time is constant. These results will provide a basis for all other constructions. The following result constructs total coupling matrices $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ on arbitrary subsets of qubits.

**Lemma 8.** *Let the coupling matrix $J$ be constant, i.e. $J_{ij} = c \in \mathbb{R}_{\geq 0}$ for all $i, j \in [n]$ with $i \neq j$. For any $\varphi \in \mathbb{R}$ and $s \in \mathbb{Z}_{\geq 0}$ the $\mathsf{GZZ}(A)$ with the matrix*

$$A = \varphi \bigoplus_{i=1}^{s} E_{k_i}, \tag{20}$$

*with $k_i \geq 1$ for $i = 1, \ldots, s$ has the encoding cost of $d = 2^{\lceil \log_2(s) \rceil} \leq 2s$ and constant total $\mathsf{GZZ}$ time $\varphi/c$. This total gate time is optimal.*

*Proof.* W.l.o.g. we set $c = \varphi = 1$. We have $n := \sum_{i=1}^{s} k_i$ qubits. Note, that $E_{k_i} = E_1 = 0$ only contributes to an entry in the diagonal of $A$, and hence this qubit does not participate in the $\mathsf{GZZ}(A)$ gate. We denote the $d \times d$ Hadamard matrix by $H^{d \times d}$ and the matrix consisting of its first $s$ columns by $H^{d \times s}$ where $s \leq n$. The orthogonality of the columns of any Hadamard matrix yields $(H^{d \times s})^T H^{d \times s} = d\,\mathbb{1}_s$. Replacing each $i$-th column by $k_i$ copies of it we obtain the $d \times n$-matrix $H_{\boldsymbol{k}}^{d \times n}$ from $H^{d \times s}$. Then, we attain the diagonal block matrix structure

$$(H_{\boldsymbol{k}}^{d \times n})^T H_{\boldsymbol{k}}^{d \times n} - d\mathbb{1}_n = d \bigoplus_{i=1}^{s} E_{k_i}. \tag{21}$$

We set the diagonal elements on the left-hand side to zero, since they only contribute to an energy offset in the Hamiltonian. Take each row of $H_{\boldsymbol{k}}^{d \times n}$ as a possible vector $\boldsymbol{m} \in \{-1, 1\}^n$ to construct the total coupling matrix

$$A = \frac{1}{d} \sum_{\boldsymbol{m} \in \mathrm{rows}(H_{\boldsymbol{k}}^{d \times n})} \boldsymbol{m}\boldsymbol{m}^T, \tag{22}$$

i.e., the time steps have been chosen as $\lambda_{\boldsymbol{m}} = \frac{1}{d}\,[\boldsymbol{m} \in \mathrm{rows}(H_{\boldsymbol{k}}^{d \times n})]$, with the Iverson bracket $[\,\cdot\,]$. Clearly, we have $\sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} = 1$. If we take the constants $c$ and $\varphi$ into account, we just multiply Eq. (22) by $\varphi/c$ which gives a total $\mathsf{GZZ}$ time of $\varphi/c$. Furthermore, this total $\mathsf{GZZ}$ time is optimal since $\|M\|_{\ell_\infty} = \|A \oslash J\|_{\ell_\infty} = \varphi/c$ satisfies the lower bound of Lemma 7. □

The encoding cost of $\mathsf{GZZ}(A)$ considered in this lemma can be reduced if redundant encodings $\boldsymbol{m} = \boldsymbol{m}' \in \{-1, 1\}^n$ are present by adding the corresponding time steps $\lambda_{\boldsymbol{m}} + \lambda_{\boldsymbol{m}'}$. It can be further reduced by using the Hadamard conjecture [20, 21]. If the Hadamard conjecture holds, then there exists Hadamard matrices of any dimension divisible by 4. It is known that the Hadamard conjecture holds for dimensions $s \leq 668$ [22, 23], thus the encoding cost can be reduced to $d = 4\lceil s/4 \rceil \leq s - 4$ in this regime. The encoding cost of the following Lemma 9, Theorems 10 and 11 and the efficient heuristic in Section 6 can be reduced in the same fashion.

The assumption in Lemma 8 of a constant coupling matrix, $J_{i \neq j} = c$, is physically unreasonable. Therefore, we relax this assumption for block sizes $k_i = 2$ which corresponds to pairwise next neighbor couplings. We use Lemma 8 to construct time-optimal $\mathsf{GZZ}$ gates for a certain family of coupling matrices.

**Lemma 9.** *Let n be even and the coupling matrix J be constant on the first subdiagonal (the other elements are arbitrary), i.e.* $J_{ij} = c$ *for* $i \in [n-1]$ *and* $j = i \pm 1$*. Then*

$$\mathsf{GZZ}(A) = \bigotimes_{i=1}^{n/2} \mathsf{GZZ}(\varphi E_2)\,, \tag{23}$$

*has the encoding cost of* $d = 2^{\lceil \log_2(n)-1 \rceil} < n$ *and constant total* GZZ *time* $\varphi/c$*. This total gate time is optimal.*

*Proof.* Since the matrix $E_2$ only has non-zero entries in the first subdiagonal, the coupling matrix $J$ only needs to be constant there. The claim follows immediately from Lemma 8 by setting $k_1 = \cdots = k_{n/2} = 2$. Clearly, the total GZZ time is optimal since it saturates the lower bound of Lemma 7. □

Lemma 9 guarantees an encoding cost of $d < n$, which is a quadratic saving compared to the general LP solution with encoding cost $\binom{n}{2}$. We note, that

$$A = \varphi \bigoplus_{i=1}^{n/2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{24}$$

corresponds to parallel $\mathsf{ZZ}(\varphi)$ gates, which find applications in simulating molecular dynamics [24]. The assumption of a constant subdiagonal of $J$ can be realized in an ion trap platform by applying an anharmonic trapping potential [16].

By combining (i), Lemma 8 and Lemma 9 we obtain the GZZ gate for next neighbor coupling.

**Theorem 10.** *Let the subdiagonal of J be constant, i.e.* $J_{ij} = c$ *for* $i \in [n-1]$ *and* $j = i \pm 1$*. Then* $\mathsf{GZZ}(A)$ *on n qubits with*

$$A = \varphi \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ & & 1 & 0 & 1 \\ & & & & \ddots \end{pmatrix}, \tag{25}$$

*has the encoding cost of* $d \leq 2n$ *(for* $n > 4$*) and constant total* GZZ *time* $2\varphi/c$*. This total gate time is optimal.*

*Proof.* We set $c = \varphi = 1$ w.l.o.g. For now, we assume that $n$ is even. Then

$$\begin{aligned} A &= \bigoplus_{i=1}^{n/2} E_2 + E_1 \bigoplus \left( \bigoplus_{i=1}^{n/2-1} E_2 \right) \bigoplus E_1 \\ &= \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{pmatrix} + \begin{pmatrix} 0 & & & & & \\ & 0 & 1 & & & \\ & 1 & 0 & & & \\ & & & 0 & 1 & \\ & & & 1 & 0 & \\ & & & & & \ddots \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ & & 1 & 0 & 1 \\ & & & & \ddots \end{pmatrix} \end{aligned} \tag{26}$$

Again, the diagonal entries do not contribute to the interactions. The first term can be implemented, using Lemma 9, with the encoding cost $d_1 = 2^{\lceil \log_2(n)-1 \rceil}$ and the GZZ time of 1. The second term can be implemented, using Lemma 8, as

$$\bigotimes_{i=1}^{s} \mathsf{GZZ}(\varphi E_{k_i})\,, \tag{27}$$

where $s = n/2+1$, $k_1 = k_s = 1$ and $k_i = 2$ for $i = 2, \ldots, n/2$, with the encoding cost $d_2 = 2^{\lceil \log_2(n+2)-1 \rceil}$ and the $\mathsf{GZZ}$ time of 1. Adding the encoding costs $d = d_1 + d_2$ and $\mathsf{GZZ}$ times of both terms yields the desired result. If $n$ is not even, then repeat the previous steps for $n+1$ but in the end reduce the dimension of all the resulting $\boldsymbol{m}$ by discarding the last entry.

This construction corresponds to a feasible solution of the primal LP (12) with the objective function value 2. To show optimality it suffices to construct a feasible solution for the dual LP (13) with the objective function value of 2. First, consider the case $n = 3$ with the total coupling matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow \boldsymbol{v}(A)^T = (1, 0, 1)\,, \tag{28}$$

and the matrix representation of the linear operator

$$V^T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}\,, \tag{29}$$

as in Eq. (10). A feasible dual solution satisfying $V^T \boldsymbol{y} \leq \boldsymbol{1}$ is $\boldsymbol{y} = (1, -1, 1)^T$. Thus, we can verify optimality for $n = 3$ since the objective function value is $\langle A, \boldsymbol{y} \rangle = \boldsymbol{v}(A)^T \boldsymbol{y} = 2$. Now, we consider arbitrary $n > 3$. Extending the dual solution for the case $n = 3$ with zeros $\boldsymbol{y} = (1, -1, 1, 0, \ldots, 0)^T$ does not change the objective function value $\langle A, \boldsymbol{y} \rangle = \boldsymbol{v}(A)^T \boldsymbol{y} = 2$. Such an extended dual solution is still feasible since $V^T$ restricted to the first three columns only has rows which are already contained in Eq. (29) due to the symmetry of Eq. (4). □

Algorithm 1 is the pseudocode implementation of Theorem 10. It takes the number of qubits $n$, the constant value $c$ of the subdiagonal of $J$ and the factor $\varphi$ of $A$ as input and returns the sparse vector $\boldsymbol{\lambda}$, containing the time steps. The encodings $\boldsymbol{m}$ are given by the indices of the non-zero elements $\lambda_{\boldsymbol{m}} \neq 0$.

**Algorithm 1** Synthesize $\mathsf{GZZ}(A)$ as in Theorem 10.

**Input:** $n, c, \varphi$
- Initialize $\boldsymbol{\lambda} = \boldsymbol{0} \in \mathbb{R}^{2^n}$
- is_odd ← $false$
- **if** $n$ odd **then**
  - $n \leftarrow n + 1$
  - is_odd ← $true$
- Let $H_1^{d_1 \times d_1}$ be a Hadamard matrix ▷ With $d_1 = 2^{\lceil \log_2(n)-1 \rceil}$ as in Lemma 9
- $H_1^{d_1 \times \frac{n}{2}} \leftarrow \frac{n}{2}$ columns of $H_1^{d_1 \times d_1}$ ▷ It does not matter which columns
- $H_1^{d_1 \times n} \leftarrow$ duplicate columns $i = 1, \ldots, \frac{n}{2}$ of $H_1^{d_1 \times \frac{n}{2}}$
- Let $H_2^{d_2 \times d_2}$ be a Hadamard matrix ▷ With $d_2 = 2^{\lceil \log_2(n+2)-1 \rceil}$
- $H_2^{d_2 \times (\frac{n}{2}+1)} \leftarrow \frac{n}{2} + 1$ columns of $H_2^{d_2 \times d_2}$ ▷ It does not matter which columns
- $H_2^{d_2 \times n} \leftarrow$ duplicate columns $i = 2, \ldots, \frac{n}{2}$ of $H_2^{d_2 \times (\frac{n}{2}+1)}$
- **for** $j \in \{1, 2\}$ **do**
  - **if** is_odd **then**
    - Delete one column of $H_j^{d_j \times n}$ ▷ It does not matter which column
  - **for** $\boldsymbol{m} \in \text{rows}(H_j^{d_j \times n})$ **do**
    - $\lambda_{\boldsymbol{m}} \leftarrow \frac{\varphi}{d_j c}$

**Output:** $\boldsymbol{\lambda}$ ▷ Can efficiently be saved in a sparse data format

The following theorem does not require any additional assumptions on $J$. It shows, how the LP (12) can be supplemented with the explicit solution to exclude certain qubits.

**Theorem 11** (Excluding qubits)**.** *Let $N$ be the total number of qubits on the quantum hardware and $n = N - s$ be the participating qubits in the* $\mathsf{GZZ}$ *gate. Synthesize the* $\mathsf{GZZ}(A)$ *gate with $A \in \mathrm{Sym}_0(\mathbb{R}^n)$, using the LP (12). Then, the total encoding cost (on $N$ qubits) is at most $\binom{n}{2}2^{\lceil \log_2(s+1)\rceil}$ and the total* $\mathsf{GZZ}$ *time, $\mathbf{1}^T\boldsymbol{\lambda}^*$, (on $N$ qubits) is the same as for the LP run on $n$ qubits.*

*Proof.* Assume w.l.o.g. that all qubits to be excluded are at the end of the qubit array. Let $k_1 = n$ and $k_2, \dots, k_{s+1} = 1$. Using Lemma 8 we obtain an encoding cost of $d_1 = 2^{\lceil \log_2(s+1)\rceil}$ and total $\mathsf{GZZ}$ time 1 to generate the matrix

$$A_1 = \begin{pmatrix} E_n & 0 \\ 0 & 0 \end{pmatrix} . \tag{30}$$

Solving the LP (12) for a matrix $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ yields the encoding cost $d_2 = \binom{n}{2}$ and the total $\mathsf{GZZ}$ time $\mathbf{1}^T\boldsymbol{\lambda}^*$. We define the extension of $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ by $A_2 \in \mathrm{Sym}_0(\mathbb{R}^N)$. The extension can be done by appending $s$ arbitrary elements in $\{-1, +1\}$ to all vectors $\boldsymbol{m} \in \{-1, +1\}^n$ given by the LP (12). Clearly,

$$A_1 \circ A_2 = \begin{pmatrix} E_n & 0 \\ 0 & 0 \end{pmatrix} \circ \begin{pmatrix} A & * \\ * & * \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} . \tag{31}$$

By (iii), the total encoding cost is $d_1 d_2$ and the total $\mathsf{GZZ}$ time is

$$\sum_{i=1}^{d_1}\sum_{j=1}^{d_2} 1/d_1 \lambda_j^* = \mathbf{1}^T\boldsymbol{\lambda}^* . \tag{32}$$

Consider now an arbitrary coupling matrix $J$. Then

$$J \circ A_1 \circ A_2 = J \circ \begin{pmatrix} E_n & 0 \\ 0 & 0 \end{pmatrix} \circ \begin{pmatrix} \tilde{A} & * \\ * & * \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} , \tag{33}$$

where $\tilde{A} = A \oslash J$ is decomposed by the LP (12). □

Algorithm 2 takes the total number of qubits $N$, the total coupling matrix $A$ (on $n$ qubits) and the set $\mathcal{Z} := \{i \in [N] \mid \text{exclude qubit } i\}$ as input and returns the sparse vector $\boldsymbol{\gamma}$, containing the time steps. The encodings $\boldsymbol{w}$ are given by the indices of the non-zero elements $\gamma_{\boldsymbol{w}} \neq 0$.

**Algorithm 2** Excluding qubits.

**Input:** $N, A, \mathcal{Z}$ (set of qubit indices to be excluded)

- Initialize $\boldsymbol{\gamma} = \mathbf{0} \in \mathbb{R}^{2^N}$
- $s \leftarrow |\mathcal{Z}|$
- $\overline{\mathcal{Z}} \leftarrow \{i \in [N] \mid i \notin \mathcal{Z}\}$
- Let $H_1^{d_1 \times d_1}$ be a Hadamard matrix ▷ With $d_1 = 2^{\lceil \log_2(s+1)\rceil}$
- $H_1^{d_1 \times (s+1)} \leftarrow s+1$ columns of $H_1^{d_1 \times d_1}$ ▷ It does not matter which columns
- $H_1^{d_1 \times N} \leftarrow$ duplicate one column of $H_1^{d_1 \times (s+1)}$ $n-1$ times and place them at indices $i \in \overline{\mathcal{Z}}$
- $\boldsymbol{\lambda}^* \leftarrow$ Solve $A = \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} \boldsymbol{m}\boldsymbol{m}^T$ ▷ using LP (12)
- **for** $\boldsymbol{v} \in \mathrm{rows}(H_1^{d_1 \times N})$ **do**
  - **for** $\boldsymbol{m} \in \{\boldsymbol{m} \mid \lambda^*_{\boldsymbol{m}} \neq 0\}$ **do** ▷ There are at most $\binom{n}{2}$ such $\boldsymbol{m}$
    - $\tilde{\boldsymbol{m}} \leftarrow$ extend $\boldsymbol{m}$ with arbitrary elements from $\{-1, +1\}$ at indices $i \in \mathcal{Z}$.
    - $\boldsymbol{w} \leftarrow \boldsymbol{v} \circ \tilde{\boldsymbol{m}}$
    - $\gamma_{\boldsymbol{w}} \leftarrow \frac{\lambda^*_{\boldsymbol{m}}}{d_1}$

**Output:** $\boldsymbol{\gamma}$ ▷ Can efficiently be saved in a sparse data format

We showed explicit constructions of time-optimal $\mathsf{GZZ}$ gates for total coupling matrices $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ with diagonal block structure and next neighbor couplings. The resulting $\mathsf{GZZ}$ gates have a constant gate time and require only linear many encodings to be implemented.

## 6 Efficient heuristic for fast GZZ gates

In this section, we build on the results of Lemma 8 to derive a heuristic algorithm for synthesizing $\mathsf{GZZ}(A)$ gates with low total gate time for any $A \in \mathrm{Sym}_0(\mathbb{R}^n)$. This algorithm runs in polynomial time as opposed to the general LP (12), which we have shown in Theorem 6 to be $\mathsf{NP}$-hard. The runtime saving is due to the restriction of the elliptope $\mathcal{E}_n$ in Eq. (16), with exponential many elements, to a set with polynomial many elements. This restriction yields a polynomial sized LP which can be solved in polynomial time. In practice, the simplex algorithm has a runtime that scales polynomial in the problem size [25].

Recall the modified Hadamard matrix $H_{\boldsymbol{k}}^{d\times n}$ defined in the proof of Lemma 8, where we used the rows of $H_{\boldsymbol{k}}^{d\times n}$ as encodings $\boldsymbol{m}$ to generate block diagonal target coupling matrices under some assumptions. Here, $s$ is the number of block matrices on the diagonal of the target matrix, $\boldsymbol{k} \in \mathbb{N}^s$ contains the dimensions for each block and $d = 2^{\lceil \log_2(s) \rceil}$ is the required number of encodings to construct such a block diagonal matrix. From now on, we only consider $\boldsymbol{k} = (j, 1 \dots 1) \in \mathbb{N}^s$ such that

$$(H_{\boldsymbol{k}}^{d\times n})^T H_{\boldsymbol{k}}^{d\times n} - d\mathbb{1}_n = d\,(E_j \oplus E_1 \oplus \cdots \oplus E_1)\ , \tag{34}$$

see Eq. (21). The requirement that $\sum_i k_i = n$ implies that such a vector $\boldsymbol{k}$ has $s = n - j + 1$ entries. Permuting the columns of $H_{\boldsymbol{k}}^{d\times n}$ results in the same permutation of the rows and columns of the right-hand side of Eq. (34). We denote the set of all column-permuted $H_{\boldsymbol{k}}^{d\times n}$ by $\mathcal{C}^{(j)}$. A specific element of $\mathcal{C}^{(j)}$ is denoted by $H_{\boldsymbol{r}}^{d\times n}$, where $\boldsymbol{r} \in \mathbb{N}^j$ is an ordered multi-index $r_1 < \cdots < r_j$ indicating which columns of $H_{\boldsymbol{r}}^{d\times n}$ are identical. For example, $\boldsymbol{r} = (2, 5, 6)$ indicates that the columns of $H_{\boldsymbol{r}}^{d\times n}$ with indices 2, 5 and 6 are identical, i.e. replacing column 5 and 6 with column 2. This notation will be useful later.

**Definition 12.** *For any $j \in \{2, 3, \dots, n\}$, we define the restricted elliptope*

$$\mathcal{E}_n^{(j)} := \left\{ \boldsymbol{m}\boldsymbol{m}^T \,\middle|\, \boldsymbol{m} \text{ is a row of } H_{\boldsymbol{r}}^{d\times n} \in \mathcal{C}^{(j)} \right\}. \tag{35}$$

*Further we define*

$$\mathcal{E}_n^{[j]} := \bigcup_{i=2}^{j} \mathcal{E}_n^{(i)}\,. \tag{36}$$

We choose the definition in Eq. (35) similar as in Eq. (16). Next, we show the size scaling of the restricted elliptopes. This directly translates to the size and runtime of the heuristic synthesis optimization.

**Proposition 13.** *For any $j \in \{2, 3, \dots, n\}$, the number of different encodings $\boldsymbol{m}$ generating the restricted elliptope $\mathcal{E}_n^{[j]}$ scales as $\mathrm{O}(n^{j+1})$.*

*Proof.* Note, that $|\mathcal{C}^{(j)}| = \binom{n}{j}$ since there are $j$ duplicate columns in $H_{\boldsymbol{r}}^{d\times n}$. The binomial coefficient can be bounded by $\binom{n}{j} \leq n^j/j!$. Since there are $d = 2^{\lceil \log_2(n-j+1) \rceil} < 2(n-j+1)$ rows of $H_{\boldsymbol{r}}^{d\times n}$ we have a rough upper bound of the number different encodings generating the restricted elliptope, $\left|\mathcal{E}_n^{(j)}\right| \leq d\binom{n}{j} < 2(n-j+1)\binom{n}{j} < 2n^{j+1}/j!$. The first inequality is due to possible redundant encodings in the definition of $\mathcal{E}_n^{(j)}$. Similarly, we can upper bound $\left|\mathcal{E}_n^{[j]}\right| \leq \sum_{i=2}^{j} \left|\mathcal{E}_n^{(i)}\right| < 2\sum_{i=2}^{j} n^{i+1}/i!$ which is a polynomial of order $j+1$. □

We denote the convex cone generated by a set $V$ by

$$\mathrm{cone}(V) := \left\{ \sum_i \lambda_i v_i \,\middle|\, \lambda_i \geq 0,\, v_i \in V \right\}. \tag{37}$$

With that, we are ready to present the main result of this section.

**Theorem 14.** $\mathrm{cone}(\mathcal{E}_n^{(2)}) = \mathrm{Sym}_0(\mathbb{R}^n)$.

*Proof.* W.l.o.g. we can assume $d = n$ and denote $H^{d\times d}_{\boldsymbol{r}} \in \mathcal{C}^{(2)}$ by $H^d_{(r_1,r_2)}$ with the property

$$(H^d_{(r_1,r_2)})^T H^d_{(r_1,r_2)} - d\mathbb{1}_d = d\, e_{(r_1,r_2)}\,, \tag{38}$$

where $e_{(r_1,r_2)}$ is an element of the standard basis for symmetric matrices with vanishing diagonal. By Eq. (38) it holds $\mathrm{Sym}_0(\mathbb{R}^n_{\geq 0}) \subseteq \mathrm{cone}(\mathcal{E}^{(2)}_n)$, i.e., symmetric matrices with non-negative entries are in the convex cone.

It is left to show that $\mathrm{Sym}_0(\mathbb{R}^n_{<0}) \subseteq \mathrm{cone}(\mathcal{E}^{(2)}_n)$, i.e., that also symmetric matrices with negative entries are in the convex cone. To show this inclusion we define $H^d_{(r_1,-r_2)}$ similar as $H^d_{(r_1,r_2)}$ except the duplicate column at $r_2$ is multiplied by $-1$ such that

$$(H^d_{(r_1,-r_2)})^T H^d_{(r_1,-r_2)} - d\mathbb{1}_d = -d\, e_{(r_1,r_2)}\,. \tag{39}$$

We have to show that for each row $\boldsymbol{m} \in \mathrm{rows}(H^d_{(r_1,-r_2)})$ there exist $\tilde{r}_1$ and $\tilde{r}_2$ such that $\boldsymbol{m} \in \mathrm{rows}(H^d_{(\tilde{r}_1,\tilde{r}_2)})$. This can be verified straightforwardly for $d = 4$ by checking all rows. W.l.o.g. we show that the hypothesis holds for any $H^{2d}_{(r_1,-r_2)}$ by assuming it holds for $H^d_{(r_1,-r_2)}$. The Sylvester-Hadamard matrix is constructed inductively according to

$$H^{2d} = \begin{pmatrix} H^d & H^d \\ H^d & -H^d \end{pmatrix}. \tag{40}$$

We consider three cases for $H^{2d}_{(r_1,-r_2)}$.

**Case 1.** For a $H^{2d}_{(r_1,-r_2)}$ with identical columns, up to minus sign, at indices $r_1, r_2 \in [d]$ or $r_1, r_2 \in [d+1, 2d]$ the hypothesis holds by our assumption by choosing $\tilde{r}_1, \tilde{r}_2 \in [d]$ or $\tilde{r}_1, \tilde{r}_2 \in [d+1, 2d]$ respectively.

**Case 2.** Considering the first $d$ rows of $H^{2d}_{(r_1,-r_2)}$ with identical columns, up to minus sign, at indices $r_1 \in [d]$ and $r_2 \in [d+1, 2d]$. This case is equivalent to *Case 1.* with $r_1, r_2 \in [d+1, 2d]$ since only the column at $r_2$ is negated.

**Case 3.** Considering the last $d$ rows of $H^{2d}_{(r_1,-r_2)}$ with identical columns, up to minus sign, at indices $r_1 \in [d]$ and $r_2 \in [d+1, 2d]$. These rows are included in the last $d$ rows of $H^{2d}_{(\tilde{r}_1,\tilde{r}_2)}$ with $\tilde{r}_1, \tilde{r}_2 \in [d+1, 2d]$ and $\tilde{r}_2 = r_2$ since the column $r_2$ is negated which is equivalent to just duplicating a column of $-H^d$.

We have shown that for each row $\boldsymbol{m} \in \mathrm{rows}(H^d_{(r_1,-r_2)})$ there exist $\tilde{r}_1$ and $\tilde{r}_2$ such that $\boldsymbol{m} \in \mathrm{rows}(H^d_{(\tilde{r}_1,\tilde{r}_2)})$. Therefore, $\mathrm{Sym}_0(\mathbb{R}^n_{<0}) \cup \mathrm{Sym}_0(\mathbb{R}^n_{\geq 0}) = \mathrm{Sym}_0(\mathbb{R}^n) = \mathrm{cone}(\mathcal{E}^{(2)}_n)$. The last equality follows from the definition of cone and $\mathcal{E}^{(2)}_n$. □

Theorem 14 shows that the constraint of LP (12) can always be fulfilled only considering $\boldsymbol{m}\boldsymbol{m}^T \in \mathcal{E}^{(2)}_n$. Similar to Eq. (10) we define the restricted linear operator $\mathcal{V}^{[j]} : \mathbb{R}^s_{\geq 0} \to \mathrm{Sym}_0(\mathbb{R}^n) : \boldsymbol{\lambda} \mapsto \sum_{\boldsymbol{m}} \lambda_{\boldsymbol{m}} \boldsymbol{m}\boldsymbol{m}^T$ for all $\boldsymbol{m}\boldsymbol{m}^T \in \mathcal{E}^{[j]}_n$ with $h = |\mathcal{E}^{[j]}_n|$, represented by a matrix $V^{[j]} \in \{-1, +1\}^{\binom{n}{2} \times h}$. We define the *restricted LP*$^j$ to be

$$\begin{aligned} \text{minimize} \quad & \mathbf{1}^T \boldsymbol{\lambda} \\ \text{subject to} \quad & V^{[j]} \boldsymbol{\lambda} = \boldsymbol{v}(M), \\ & \boldsymbol{\lambda} \in \mathbb{R}^h_{\geq 0}\,. \end{aligned} \tag{41}$$

Algorithm 3 summarizes the steps to construct $\mathcal{E}^{[j]}_n$ and therefore the matrix representation of the restricted linear operator $V^{[j]}$. In practice, Algorithm 3 has to be executed only once per number of qubits $n$ since the constraints of LP (41) can be fulfilled for all $M \in \mathrm{Sym}_0(\mathbb{R}^n)$. This is due to the fact that $\mathcal{E}^{(2)}_n \subseteq \mathcal{E}^{[j]}_n$ for any $2 \leq j \leq n$ and $\mathrm{cone}(\mathcal{E}^{(2)}_n) = \mathrm{Sym}_0(\mathbb{R}^n)$ (Theorem 14). The time and space complexity of Algorithm 3 scales polynomially in $n$ as shown in Proposition 13.

Therefore, the restricted LP$^j$ is also of polynomial size for a fixed $j$. Increasing $j$ leads to better approximations due to the enlarged search space for the optimal solution. Note, that the runtime of the mixed integer program (MIP) defined in [6, Section 2.2.2] also benefits from using $\mathcal{E}_n^{[j]}$.

As mentioned in Section 5 the dimension of the Hadamard matrices in Eq. (34) can be reduced to $d = 4\lceil (n-1)/4 \rceil \leq n-5$ if we assume that the Hadamard conjecture holds. Therefore, the runtime of the restricted LP$^j$ is reduced as well if such Hadamard matrices are used. In Section 8 we numerically benchmark the heuristic algorithm with and without the reduced runtime of the restricted LP$^j$.

**Algorithm 3** Constructing $\mathcal{E}_n^{[j]}$.

**Input:** $n, j$

Initialize $\mathcal{E}_n^{[j]} = \emptyset$
**for** $i = 2, \ldots, j$ **do**
  Initialize $\mathcal{E}_n^{(i)} = \emptyset$
  Let $H^{d\times d}$ be a Hadamard matrix ▷ With $d = 2^{\lceil \log_2(n-i+1) \rceil}$
  $H^{d\times(n-i+1)} \leftarrow n-i+1$ columns of $H^{d\times d}$ ▷ It does not matter which columns
  **for** all $\boldsymbol{r}$ such that $1 \leq r_1 < \cdots < r_i \leq n$ **do** ▷ There are $\binom{n}{i}$ such $\boldsymbol{r}$
    $H_{\boldsymbol{r}}^{d\times n} \leftarrow$ duplicate one column of $H^{d\times(n-i+1)}$ $i-1$ times to indices $r_1, \ldots, r_i$
    $\mathcal{E}_n^{(i)} \leftarrow \mathcal{E}_n^{(i)} \cup \left\{ \boldsymbol{m}\boldsymbol{m}^T \,\middle|\, \boldsymbol{m} \text{ is a row of } H_{\boldsymbol{r}}^{d\times n} \right\}$
  $\mathcal{E}_n^{[j]} \leftarrow \mathcal{E}_n^{[j]} \cup \mathcal{E}_n^{(i)}$

**Output:** $\mathcal{E}_n^{[j]}$

## 7 Bounds on the total GZZ time

Our main analytic result (Theorem 15) is that the optimal total GZZ time $\mathbf{1}^T\boldsymbol{\lambda}^*$ is lower and upper bounded by the norms $\|M\|_{\ell_\infty}$ and $\|M\|_{\ell_1}$, respectively. Note, that for a dense matrix $M$, its norm $\|M\|_{\ell_1}$ scales quadratic with the number of qubits $n$. We conjecture an improved upper bound on the total GZZ time for dense $M$ that scales at most linear with the number of qubits $n$. We support this conjecture with explicit solutions for the LP (12) reaching this bound for any $n$ and numerical results validating the conjecture for $n \leq 8$.

**Theorem 15.** *The optimal total gate time of* GZZ$(A)$ *with* $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ *is lower and upper bounded by*

$$\|M\|_{\ell_\infty} \leq \mathbf{1}^T\boldsymbol{\lambda}^* \leq \|M\|_{\ell_1}\,, \tag{42}$$

*where* $M := A \oslash J$. *Equality holds for the lower bound for all matrices* $M = C\boldsymbol{m}\boldsymbol{m}^T$ *for any* $\boldsymbol{m} \in \{-1,+1\}^n$ *and* $C \geq 0$.

*Proof.* The lower bound has been shown in Lemma 7. Equality in the lower bound holds for $M = C\boldsymbol{m}\boldsymbol{m}^T$ by setting $\lambda_{\boldsymbol{m}} = C = \|M\|_{\ell_\infty}$ and $\lambda_{\boldsymbol{m}'} = 0$ for all $\boldsymbol{m}' \neq \boldsymbol{m}$. We use the explicit construction of the standard basis elements for symmetric matrices from the proof of Theorem 14 to show the upper bound. To be precise, we have

$$\frac{|M_{ij}|}{d}(H_{\boldsymbol{r}}^{d\times n})^T H_{\boldsymbol{r}}^{d\times n} - d\mathbb{1}_n = M_{ij} e_{(i,j)}\,, \tag{43}$$

where $\boldsymbol{r} = (i,j)$ if $M_{ij} \geq 0$ or $\boldsymbol{r} = (\tilde{i},\tilde{j})$ if $M_{ij} < 0$ as in Eqs. (38) and (39), respectively. We define $\boldsymbol{\lambda}^{(i,j)}$ with the entries $\lambda_{\boldsymbol{m}}^{(i,j)} := \frac{|M_{ij}|}{d}[\boldsymbol{m} \in \mathrm{rows}(H_{\boldsymbol{r}}^d)]$. According to Lemma 8 we have $\mathbf{1}^T\boldsymbol{\lambda}^{(i,j)} = |M_{ij}|$. Adding $\boldsymbol{\lambda}^{(i,j)}$ for all $i < j$ yields the upper bound $\|M\|_{\ell_1}$. □

These bounds get tighter the sparser $M$ is. If $M$ has only one non-zero value, then clearly $\|M\|_{\ell_\infty} = \mathbf{1}^T\boldsymbol{\lambda}^* = \|M\|_{\ell_1}$. Furthermore, these bounds also hold for the heuristic, which we presented in Section 6.

Next, we state our conjecture that the optimal gate time of GZZ$(A)$ scales at most linear with the number of qubits.

**Conjecture 16.** *The optimal gate time of* $\mathsf{GZZ}(A)$ *with* $A \in \mathrm{Sym}_0(\mathbb{R}^n)$ *is tightly upper bounded by*

$$\mathbf{1}^T \boldsymbol{\lambda}^* \leq \|M\|_{\ell_\infty} \cdot \begin{cases} n\,, & \textit{for odd } n\,, \\ n-1\,, & \textit{for even } n\,. \end{cases} \tag{44}$$

Hence, it provides a tighter bound for dense $M$ than Theorem 15. To support the claim of Conjecture 16 we first construct explicit dual and primal feasible solutions for the case $M = -E_n$ for the LP's (12) and (13), respectively. Then optimality is given by showing equality of the objective function values of the primal and dual problem. We further show that the case $M = -E_n$ leads to the same objective function value as $M = -\boldsymbol{m}\boldsymbol{m}^T$ for any $\boldsymbol{m} \in \{-1, 1\}^n$. Finally, we provide numerical evidence that the conjecture holds for $n \leq 8$.

For practical purposes it is important to keep in mind that the platform given $J$ matrix might also scale with the number of qubits resulting in a qubit dependent constant $\|A \oslash J_n\|_{\ell_\infty} = \|M_n\|_{\ell_\infty}$.

## 7.1 Explicit solutions for $M = -\boldsymbol{m}\boldsymbol{m}^T$

The following lemma will be used in the proof of the explicit feasible solution of the dual problem for $M$ being of the form $M = -\boldsymbol{m}\boldsymbol{m}^T$. We can identify $\boldsymbol{m} \in \{-1, 1\}^n$ with $\boldsymbol{b} \in \mathbb{F}_2^n$ via $\boldsymbol{m} = (-1)^{\boldsymbol{b}}$ as explained in Section 2.

**Lemma 17.** *It holds that*

$$P_{\boldsymbol{b}} := \sum_{i<j}^{n} (-1)^{b_i \oplus b_j} = \binom{n}{2} - 2|\boldsymbol{b}|(n - |\boldsymbol{b}|)\,, \tag{45}$$

*for any binary vector* $\boldsymbol{b} \in \mathbb{F}_2^n$. *We denote the Hamming weight by* $|\boldsymbol{b}|$.

*Proof.* Let $\boldsymbol{m} = (-1)^{\boldsymbol{b}}$. If the Hamming weight $|\boldsymbol{b}|$ vanishes, then $\boldsymbol{m} = (+1, \ldots, +1)$ and $P_{\boldsymbol{b}} = \binom{n}{2}$ which is the maximal value. If $|\boldsymbol{b}| \neq 0$, $\boldsymbol{m}$ contains $|\boldsymbol{b}|$ entries $-1$, such that the upper triangular part of $\boldsymbol{m}\boldsymbol{m}^T$ contains a rectangle of $-1$'s with length $|\boldsymbol{b}|$ and width $n - |\boldsymbol{b}|$ so the total amount of $-1$'s is $|\boldsymbol{b}|(n - |\boldsymbol{b}|)$. Therefore,

$$P_{\boldsymbol{b}} = \binom{n}{2} - 2|\boldsymbol{b}|(n - |\boldsymbol{b}|)\,. \tag{46}$$

□

**Lemma 18** (explicit dual feasible solution)**.** *Let* $M = -E_n$*, then there is an explicit feasible solution* $\boldsymbol{y}$ *to the dual LP (13) with*

$$\langle -E_n, \boldsymbol{y} \rangle = \begin{cases} n\,, & \textit{for odd } n\,, \\ n-1\,, & \textit{for even } n\,. \end{cases} \tag{47}$$

*Proof.* We assume that $y = y_1 = y_2 = \ldots$. Therefore, it suffices to show that

$$y \sum_{i<j}^{n} (-1)^{b_i \oplus b_j} \leq 1\,. \tag{48}$$

From Lemma 17 we know that $\boldsymbol{y} = 1/\min(P_{\boldsymbol{b}})\mathbf{1}$ satisfies the constraint in Eq. (48). The minimum $\min(P_{\boldsymbol{b}}) = -\lfloor n/2 \rfloor$ is reached for $|\boldsymbol{b}| = \lceil n/2 \rceil$ or $|\boldsymbol{b}| = \lfloor n/2 \rfloor$ which can be verified from the expression of $P_{\boldsymbol{b}}$ in Lemma 17. Thus, we obtain $\boldsymbol{y} = -1/\lfloor n/2 \rfloor \mathbf{1}$. The objective function evaluates to

$$\langle -E_n, \boldsymbol{y} \rangle = -\mathbf{1}^T \boldsymbol{y} = \frac{\binom{n}{2}}{\lfloor n/2 \rfloor} = \begin{cases} n\,, & \text{for odd } n\,, \\ n-1\,, & \text{for even } n\,. \end{cases} \tag{49}$$

□

For the construction of a feasible solution to the primal problem we first require the following result.

**Lemma 19.** *Let $k < n$ be natural numbers. Then*

$$\sum_{\boldsymbol{b}\in\mathbb{F}_2^n:|\boldsymbol{b}|=k} |b_i \oplus b_j| = 2\binom{n-2}{k-1} \tag{50}$$

*for all $i, j \in [n]$ with $i \neq j$.*

*Proof.* Consider the case $n = 4$ and $k = 2$. We get $\binom{4}{2} = 6$ binary vectors with $|\boldsymbol{b}| = 2$. It can be easily verified that $\sum_{|\boldsymbol{b}|=2} |b_i \oplus b_j| = 4 = 2\binom{2}{1}$. Now, we assume that for a given $n$ and $k < n$ the Eq. (50) holds. It suffices to verify Eq. (50) for $i \leq n$ and $j = n + 1$. We fix $k$, define $n' := n + 1$ and take a $\boldsymbol{b} \in \mathbb{F}_2^{n'}$. We have $\binom{n+1}{k}$ binary vectors with $|\boldsymbol{b}| = k$.

For the case $b_{n+1} = 0$ we have $\binom{n}{k}$ such vectors and

$$\sum_{\substack{|\boldsymbol{b}|=k\\ b_{n+1}=0}} |b_i \oplus b_{n+1}| = \sum_{|\boldsymbol{b}|=k} |b_i| = \binom{n-1}{k-1}, \tag{51}$$

for all $i \leq n$. There are $\binom{n-1}{k-1}$ different ways in placing $k - 1$ 1's.

For the case $b_{n+1} = 1$ we have $\binom{n}{k-1}$ such vectors and

$$\sum_{\substack{|\boldsymbol{b}|=k\\ b_{n+1}=1}} |b_i \oplus b_{n+1}| = \sum_{|\boldsymbol{b}|=k} |b_i \oplus 1| = \binom{n-1}{k-1}, \tag{52}$$

for all $i \leq n$. There are $\binom{n-1}{k-1}$ different ways in placing $k - 1$ 0's.

Combining the two cases $b_{n+1} = 0$ and $b_{n+1} = 1$ we obtain

$$\sum_{|\boldsymbol{b}|=k} |b_i \oplus 0| + |b_i \oplus 1| = 2\binom{n-1}{k-1} = 2\binom{n'-2}{k-1}, \tag{53}$$

for all $i \leq n$. □

We motivate the next lemma with the result of the explicit dual feasible solution for $M = -E_n$ from Lemma 18 and the complementary slackness condition. The complementary slackness condition for a linear program states that if the $i$-th inequality of the dual problem is a strict inequality for a feasible solution $\boldsymbol{y}$, then the $i$-th component of a feasible solution of the primal problem $\boldsymbol{\lambda}$ is zero:

$$(V^T\boldsymbol{y})_i < 1 \Rightarrow \lambda_i = 0\,. \tag{54}$$

We use this in the following lemma to construct a feasible solution for the primal LP (12).

**Lemma 20** (explicit primal feasible solution)**.** *Let $M = -E_n$, then there is an explicit feasible solution $\boldsymbol{\lambda}$ to the primal LP (12) with*

$$\mathbf{1}^T\boldsymbol{\lambda} = \begin{cases} n\,, & \textit{for odd } n\,,\\ n-1\,, & \textit{for even } n\,. \end{cases} \tag{55}$$

*Proof.* For this proof we define

$$k := \begin{cases} n\,, & \text{for odd } n\,,\\ n-1\,, & \text{for even } n\,. \end{cases} \tag{56}$$

We only consider binary vectors of the set $S := \{\boldsymbol{b} \in \mathbb{F}_2^n \mid |\boldsymbol{b}| = \lceil n/2\rceil, \lfloor n/2\rfloor \text{ and } b_n = 0\}$. This set is motivated by the complementary slackness condition and Lemma 18. It can be calculated that $|S| = \binom{k}{\lfloor k/2\rfloor}$ using the recurrence relation

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \tag{57}$$

for the binomial coefficients. We show that

$$\sum_{\boldsymbol{b}\in S}(-1)^{b_i\oplus b_j} = -D_n\,, \tag{58}$$

for a constant $D_n > 0$, which we calculate later. If not explicitly stated, all equations in this proof containing $i$, $j$ hold for all $i, j \in [n]$, $i \neq j$. We denote $\boldsymbol{\lambda}_S$ by all $\lambda_{\boldsymbol{m}} = \lambda_{\boldsymbol{b}}$ corresponding to the encoding $\boldsymbol{m} = (-1)^{\boldsymbol{b}}$ with $\boldsymbol{b} \in S$. If Eq. (58) holds, we can choose a $\boldsymbol{\lambda}_S = 1/D_n\mathbf{1}$ resulting in

$$\sum_{\boldsymbol{b}\in S}\boldsymbol{\lambda}_{\boldsymbol{b}}(-1)^{b_i\oplus b_j} = -1\,, \tag{59}$$

which implies feasibility for $M = -E_n$. It is left to show Eq. (58) and determine $D_n$.

First, we consider odd $n$. By definition of $S$ we have that $b_n = 0$ for all $\boldsymbol{b}$. Therefore, we obtain $\binom{n-1}{\lfloor n/2\rfloor} + \binom{n-1}{\lceil n/2\rceil}$ binary vectors with $|\boldsymbol{b}| = \lceil n/2\rceil$ or $|\boldsymbol{b}| = \lfloor n/2\rfloor$ respectively. Counting the occurrences of "$-1$" in the sum of Eq. (58) is equivalent to counting the occurrences of "1" in the sum

$$\begin{aligned}\sum_{\boldsymbol{b}\in S}|b_i\oplus b_j| &= 2\left(\binom{n-3}{\lfloor\frac{n}{2}\rfloor-1} + \binom{n-3}{\lceil\frac{n}{2}\rceil-1}\right)\\ &= 2\binom{n-2}{\lfloor\frac{n}{2}\rfloor}\,,\end{aligned} \tag{60}$$

where we used Lemma 19, the recurrence relation for the binomial coefficients and $\lceil n/2\rceil - 1 = \lfloor n/2\rfloor$. Counting the occurrences of "$+1$" in the sum of Eq. (58) yields

$$\binom{n-1}{\lfloor\frac{n}{2}\rfloor} + \binom{n-1}{\lceil\frac{n}{2}\rceil} - 2\binom{n-2}{\lfloor\frac{n}{2}\rfloor} = \binom{n}{\lfloor\frac{n}{2}\rfloor} - 2\binom{n-2}{\lfloor\frac{n}{2}\rfloor}\,. \tag{61}$$

where we used $\binom{n}{\lfloor n/2\rfloor+1} = \binom{n}{\lfloor n/2\rfloor}$ for odd $n$. We now evaluate Eq. (58) for odd $n$

$$\begin{aligned}-D_n = \sum_{\boldsymbol{b}\in S}(-1)^{b_i\oplus b_j} &= \binom{n}{\lfloor\frac{n}{2}\rfloor} - 4\binom{n-2}{\lfloor\frac{n}{2}\rfloor}\\ &= \binom{n}{\lfloor\frac{n}{2}\rfloor} - \frac{n+1}{n}\binom{n}{\lfloor\frac{n}{2}\rfloor}\\ &= \frac{-1}{n}\binom{n}{\lfloor\frac{n}{2}\rfloor}\,.\end{aligned} \tag{62}$$

The case for even $n$ follows the same steps as for odd $n$, resulting in

$$\begin{aligned}-D_n = \sum_{|\boldsymbol{b}|=\frac{n}{2}}(-1)^{b_i\oplus b_j} &= \binom{n-1}{\frac{n}{2}} - \frac{n}{n-1}\binom{n-1}{\frac{n}{2}}\\ &= \frac{-1}{n-1}\binom{n-1}{\frac{n}{2}}\,.\end{aligned} \tag{63}$$

Equations (62) and (63) show that $D_n = 1/k\binom{k}{\lfloor k/2\rfloor}$ and that Eq. (58) holds. Since $|S| = \binom{k}{\lfloor k/2\rfloor}$ the objective function value is

$$\mathbf{1}^T\boldsymbol{\lambda}_S = \binom{k}{\lfloor k/2\rfloor}\frac{1}{\frac{1}{k}\binom{k}{\lfloor k/2\rfloor}} = k\,. \tag{64}$$

□

From the equality of the objective functions for the primal and dual problem from Lemma 20 and Lemma 18 respectively we know that the proposed dual/primal feasible solutions for $M = -E_n$ are in fact optimal solutions. Now, we show that the GZZ gate time with $M = -\boldsymbol{m}\boldsymbol{m}^T$ for any $\boldsymbol{m} \in \{-1, +1\}^n$ is the same as for the case $M = -E_n$.

**Theorem 21.** *If* $A \oslash J =: M = -C(\boldsymbol{m}\boldsymbol{m}^T)$ *for any* $\boldsymbol{m} \in \{-1,+1\}^n$ *and* $C \geq 0$*, then the optimal gate time of* GZZ$(A)$ *is*

$$\mathbf{1}^T \boldsymbol{\lambda}^* = \|M\|_{\ell_\infty} \begin{cases} n\,, & \text{for odd } n\,, \\ n-1\,, & \text{for even } n\,. \end{cases} \tag{65}$$

*Proof.* The statement has been shown for the case $M = -E_n$ by constructing an explicit solution. It is left to show that the cases $M = -C(\boldsymbol{m}\boldsymbol{m}^T)$ for a constant $C \geq 0$ yield the same objective function value. Since the objective function and the constraints are linear we can w.l.o.g. assume $C = 1$. It is clear, that $\operatorname{sign}(M_{ij}) = \operatorname{sign}(-m_i m_j) = -(-1)^{c_i \oplus c_j}$ for all $i < j$, with $\boldsymbol{m} = (-1)^{\boldsymbol{c}}$. Let $\boldsymbol{y} = -y \operatorname{sign}(\boldsymbol{v}(M))$ for a $y \in \mathbb{R}$, then each constraint of $V^T \boldsymbol{y} \leq \mathbf{1}$ of the dual LP (13) reads as

$$y \sum_{i<j} (-1)^{b_i \oplus b_j \oplus c_i \oplus c_j} = y \sum_{i<j} (-1)^{\tilde{b}_i \oplus \tilde{b}_j} \leq 1\,, \tag{66}$$

with $\tilde{\boldsymbol{b}} \coloneqq \boldsymbol{b} \oplus \boldsymbol{c}$ element wise. Consider the ordered set of all $\boldsymbol{b} \in \mathbb{F}_2^n$ with $b_n = 0$, then $\tilde{\boldsymbol{b}}$ is just a permutation of that set. Due to the permutation symmetry of the qubits the optimal value of the LP (12) for any $M = -(\boldsymbol{m}\boldsymbol{m}^T)$ is the same as for the case $M = -E_n$. Setting $\boldsymbol{y} = -y \operatorname{sign}(\boldsymbol{v}(M))$ with $y = 1/\lfloor n/2 \rfloor$ as in the proof of Lemma 18 yields an optimal solution to the dual LP (13) for the case $M = -(\boldsymbol{m}\boldsymbol{m}^T)$. □

Note that, trivially, the lower bound $\mathbf{1}^T \boldsymbol{\lambda}^* = \|M\|_{\ell_\infty}$ is reached if $M = C(\boldsymbol{m}\boldsymbol{m}^T)$ for any $\boldsymbol{m} \in \{-1,+1\}^n$ and $C \geq 0$.

One possibility to prove Conjecture 16 is to show, that the matrix $M = -E_n$ maximizes the value of the LP (12) among all matrices $M \in \operatorname{Sym}_0([-1,+1]^n)$. To this end, consider the LP

$$\begin{aligned} \text{maximize} \quad & \|\boldsymbol{y}\|_{\ell_1} \\ \text{subject to} \quad & V^T \boldsymbol{y} \leq \mathbf{1}, \\ & \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}}\,, \end{aligned} \tag{67}$$

which is independent of $M \in \operatorname{Sym}_0([-1,+1]^n)$. It holds that

$$\max_{V^T \boldsymbol{y} \leq \mathbf{1}} \left( \max_{\|M\|_{\ell_\infty} \leq 1} \langle M, \boldsymbol{y} \rangle \right) = \max_{V^T \boldsymbol{y} \leq \mathbf{1}} \|\boldsymbol{y}\|_{\ell_1} \tag{68}$$

according to $\|\boldsymbol{x}\|_{\ell_1} = \max_{\|\boldsymbol{p}\|_{\ell_\infty} \leq 1} \boldsymbol{p}^T \boldsymbol{x}$. Therefore, the optimal objective value of LP (67) is an upper bound on all optimal objective values of LP (12). Clearly, the constructed solution in Lemma 18 is feasible for LP (67). Unfortunately, proving that this constructed solution is optimal is quite challenging, as we discuss in Appendix A.

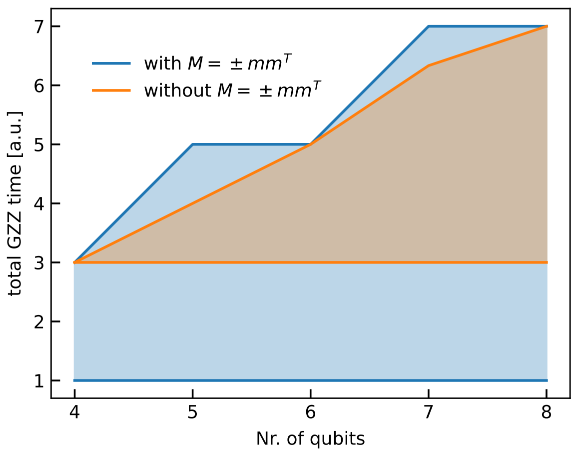


Figure 1: The optimal objective function value of the dual LP (13) over the number of qubits. **Blue:** The range of optimal values for all binary $M \in \operatorname{Sym}_0(\{-1,+1\}^n)$. **Orange:** The range of optimal values for all binary $M \in \operatorname{Sym}_0(\{-1,+1\}^n)$ without $M = \pm(\boldsymbol{m}\boldsymbol{m}^T)$ for any $\boldsymbol{m} \in \{-1,+1\}^n$.

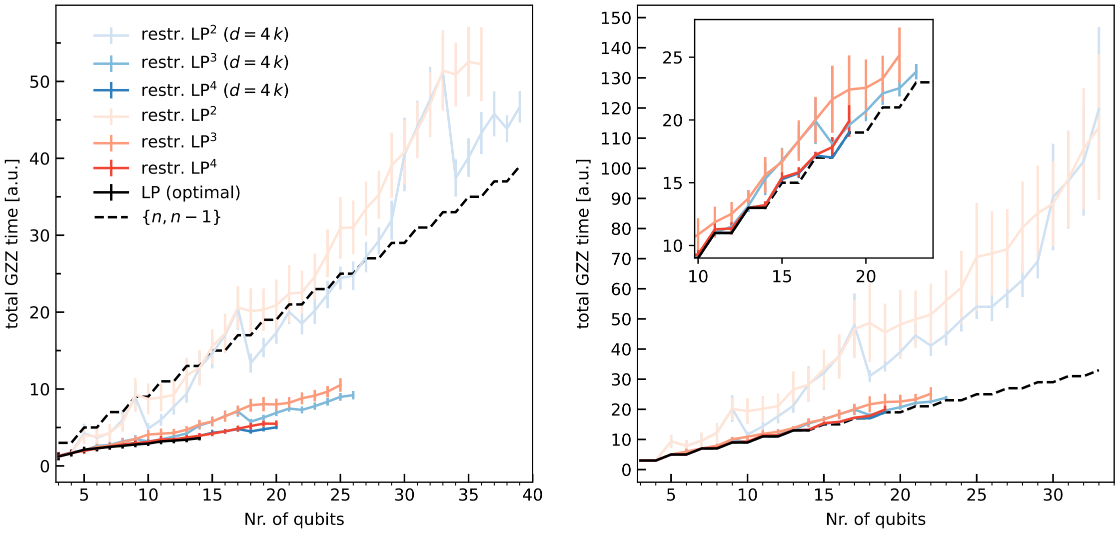


Figure 2: Comparing the performance of the original (optimal) LP (12) and the restricted LP$^j$ (41) for $j = 2, 3, 4$. For each line we let the LP's run for a fixed time. The black dashed line is the upper bound for the original LP (12) from Conjecture 16. The reddish lines show the total GZZ times for the restricted LP$^j$ (41) for $j = 2, 3, 4$. The blueish lines show the total GZZ times for the restricted LP$^j$ (41) using Hadamard matrices of dimension $d = 4\,k$ to construct the restricted linear operator $V^{[j]}$. **Left:** Average case scaling of the total GZZ times. The data points and error bars show the mean and the standard deviation over 100 uniformly sampled matrices $A_{ij} = A_{ji} \in [-1, 1]$ for $i < j$. **Right:** Conjectured worst-case scaling of the total GZZ times. The data points and error bars show the mean and the standard deviation over 100 binary matrices $A = -\boldsymbol{m}\boldsymbol{m}^T$ with uniformly sampled encodings $\boldsymbol{m} \in \{-1, 1\}^n$.

## 8 Numerical results

### 8.1 Numeric validation of Conjecture 16 for small $n$

For the numeric validation of the conjecture for small $n$ we solve the dual LP (13) for all binary $M \in \mathrm{Sym}_0(\{-1, +1\}^n)$ of which there are $2^{\binom{n}{2}}$. For $n = 3$ there are only binary matrices of the form $M = \pm(\boldsymbol{m}\boldsymbol{m}^T)$ and by Theorem 21 the conjecture holds. Figure 1 shows that Conjecture 16 holds for $n \leq 8$. For odd $n \leq 8$ the cases $M = -(\boldsymbol{m}\boldsymbol{m}^T)$ are in fact the only cases reaching the upper bound. This can be seen in Fig. 1 by the blue area exceeding the orange area, which only consists of the optimal values for all binary matrices without $M = \pm(\boldsymbol{m}\boldsymbol{m}^T)$ for any $\boldsymbol{m} \in \{-1, +1\}^n$.

### 8.2 Numerical benchmark for the heuristic

We compare the performance of the restricted LP$^j$ (41) to the original LP (12). To this end we provide numerical results on the average-case and the conjectured worst-case scaling of the total GZZ time. The left plot in Fig. 2 shows the average-case scaling of the total GZZ time. GZZ$(A)$ gates with uniformly sampled matrix elements $A_{ij} = A_{ji} \in [-1, 1]$ for all $i < j$ and $i, j \in [n]$ are synthesized. For the worst-case scaling of the right plot in Fig. 2, GZZ$(A)$ gates with binary matrices $A = -\boldsymbol{m}\boldsymbol{m}^T$ with uniformly sampled encodings $\boldsymbol{m} \in \{-1, 1\}^n$ are synthesized. We note that these are the matrices with the conjectured worst-case scaling for the LP and that the LP$^j$ might have different worst-case matrices $A$. For convenience, we have set $M = A$, i.e. setting $J = E_n$ which omits the hardware specific time units given by the quantum platform. For realistic $J$'s and GZZ$(A)$ gate times we refer to [6]. The Python package CVXPY [26, 27] with the GNU linear program kit simplex solver [28] is used to solve the LP (12) and the restricted LP$^j$ (41).

Figure 2 shows the total GZZ time over the number of qubits $n$ for the restricted LP$^j$ (41) and LP (12). We assigned a fixed runtime (20 minutes) for each LP to synthesize GZZ$(A)$ gates for all sample matrices $A$ and as many $n$ as possible. Clearly, the runtime of the heuristic algorithm is much shorter than the runtime of the optimal LP (12). Increasing the hierarchy of the heuristic

$j > 2$ reduces the total GZZ time significantly while still maintaining a short runtime. The total GZZ time obtained from the heuristic also seems to scale linear with the number of qubits although with a different scaling constant. As mentioned in Sections 5 and 6 the size of the restricted $\mathrm{LP}^j$ (41), and therefore the runtime, can be further reduced. This reduction is achieved by using Hadamard matrices of dimension $d = 4\,k$ with $k \in \mathbb{N}$ instead of Sylvester-Hadamard matrices of dimension $d = 2^k$ in the construction of the restricted linear operator $V^{[j]}$. This is based on the Hadamard conjecture, which is known to be true for $d \leq 668$ [22, 23]. Surprisingly, using these Hadamard matrices with $d = 4\,k$ not only yields shorter runtime of the heuristic algorithm but also a significant reduction of the total GZZ time compared to the original restricted $\mathrm{LP}^j$ (41).

Our numerical results show that the heuristic algorithm approximates well the optimal total GZZ time, while maintaining a short runtime. This holds true for both, the average-case and the conjectured worst-case scaling. Therefore, we hope that this heuristic will prove to be an important tool to implement fast GZZ gates in practice.

## 9 Conclusion

We investigated the time-optimal multi-qubit gate synthesis introduced in Ref. [6]. We show that synthesizing time-optimal multi-qubit gates in our setting is NP-hard. However, we also provide explicit solutions for certain cases with constant gate time and a polynomial-time heuristics to synthesize fast multi-qubit gates. Our numerical simulations suggest that these heuristics provide good approximations to the optimal GZZ gate time. Furthermore, tight bounds on the scaling of the optimal multi-qubit gate times were shown. More precisely, we showed that the optimal multi-qubit gate time scales at most as $\|A \oslash J\|_{\ell_1}$, the $\ell_1$-norm of the element-wise division of the total and physical coupling matrices $A$ and $J$, respectively. We also conjectured that the optimal GZZ gate time scales at most linear with the number of qubits. Our results are practical to estimate the execution time of a given circuit, where the entangling gates are implemented as GZZ gates. The execution time is a crucial parameter, in particular, in the NISQ era since it is limiting the length of a gate sequence due to finite coherence time.

It is our hope to proof the conjectured linear scaling of the optimal GZZ gate time in the near future. Moreover, we would like to test and verify our proposed time-optimal multi-qubit gate synthesis methods in an experiment. Depending on the quantum platform we would like to develop adapted error mitigation schemes for the GZZ gates and investigate their robustness against errors.

## Acknowledgements

We are grateful to Lennart Bittel and Arne Heimendahl for fruitful discussions on complexity theory and convex optimization, respectively. We also want to thank Frank Vallentin for valuable comments on our conjecture and proof ideas.

This work has been funded by the German Federal Ministry of Education and Research (BMBF) within the funding program “quantum technologies – from basic research to market” via the joint project MIQRO (grant number 13N15522) and the Fujitsu Services GmbH as part of an endowed professorship “Quantum Inspired and Quantum Optimization”.

# Appendices

## A Challenges in proving Conjecture 16

In this section, we want to discuss some obstacles encountered trying to proof Conjecture 16. Conjecture 16 holds, if we show that our constructed solutions from Section 7.1 are optimal solutions for LP (67). Meaning that there is no other feasible solution resulting in a larger objective function value compared to our constructed solution.

We tried an inductive proof which turned out to be intricate due to the additional degrees of freedom in each induction step. Furthermore, we utilized the connection to graph theory from Lemma 2 to transform the LP (67) to a LP over the cut polytope. There, the challenge is the affine mapping between the elliptope and the cut polytope in Eq. (16) and Definition 1, which alters the optimization problem crucially. In the following, we discuss another approach based on showing the sufficiency of the Karush–Kuhn–Tucker (KKT) conditions in more detail.

### A.1 Concave program

A convex linear program in standard form minimizes a convex objective function over a convex set. But LP (67) *maximizes* a convex objective function over a convex set. Such optimizations are called *concave programs*. It is known that the maximum is attained at the extreme points of the polytope $V^T\boldsymbol{y} \leq \mathbf{1}$ and therefore might have many local optima [29]. There are several equivalent sufficient conditions for optimality [30]. Here, we investigate one in detail. First, we define the *conjugate function* for a function $f : \mathbb{R}^n \to \mathbb{R}$

$$f^*(\boldsymbol{x}) \coloneqq \sup\big\{\boldsymbol{x}^T\boldsymbol{y} - f(\boldsymbol{y}) \,\big|\, \boldsymbol{y} \in \mathcal{D}(f)\big\}\,, \tag{69}$$

with $\mathcal{D}(f)$ the domain of $f$. Furthermore, we need the *support function* $h_A : \mathbb{R}^n \to \mathbb{R}$ for a closed convex set $A$

$$h_A(\boldsymbol{x}) \coloneqq \sup\big\{\boldsymbol{x}^T\boldsymbol{y} \,\big|\, \boldsymbol{y} \in A\big\}\,. \tag{70}$$

Then the sufficient optimality condition in our case is

$$\|\boldsymbol{y}^*\|_{\ell_1} = \sup\Big\{h_A(\boldsymbol{x}) - (\|\boldsymbol{x}\|_{\ell_1})^* \,\Big|\, \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}}\Big\}\,, \tag{71}$$

with $A = \Big\{\boldsymbol{x} \in \mathbb{R}^{\binom{n}{2}} \,\Big|\, V^T\boldsymbol{x} \leq \mathbf{1}\Big\}$. The conjugate function of $\|\boldsymbol{x}\|_{\ell_1}$ is

$$(\|\boldsymbol{x}\|_{\ell_1})^* = \begin{cases} 0\,, & \text{if } \|\boldsymbol{x}\|_{\ell_\infty} \leq 1\,,\\ \infty\,, & \text{otherwise}\,.\end{cases} \tag{72}$$

Then we have

$$\begin{aligned}\|\boldsymbol{y}^*\|_{\ell_1} &= \sup\{h_A(\boldsymbol{y}) \mid \|\boldsymbol{y}\|_{\ell_\infty} \leq 1\}\\ &= \sup\big\{\boldsymbol{x}^T\boldsymbol{y} \,\big|\, \|\boldsymbol{y}\|_{\ell_\infty} \leq 1, V^T\boldsymbol{x} \leq \mathbf{1}\big\}\\ &= \sup\big\{\|\boldsymbol{x}\|_{\ell_1} \,\big|\, V^T\boldsymbol{x} \leq \mathbf{1}\big\}\,,\end{aligned} \tag{73}$$

which is the same formulation as the original LP (67).

### A.2 Dualization

If we can formulate the dual LP to the primal LP (67) and find a feasible solution, then the dual objective function value upper bounds the primal objective function value by weak duality [31]. The standard form of an optimization problem with only linear inequality constraints is

$$\begin{aligned}\text{minimize}\quad & f(\boldsymbol{y})\\ \text{subject to}\quad & A\boldsymbol{y} \leq \boldsymbol{b}.\end{aligned} \tag{74}$$

Then, the Lagrange dual function is given by

$$g(\boldsymbol{\lambda}) = -\boldsymbol{b}^T\boldsymbol{\lambda} - f^*(-A^T\boldsymbol{\lambda})\,, \tag{75}$$

where $f^*$ denotes the conjugate function [14]. For LP (67) we have $f(\boldsymbol{y}) = -\|\boldsymbol{y}\|_{\ell_1}$ (minus sign due to the minimization in the standard optimization form).

$$
\begin{aligned}
(-\|\boldsymbol{x}\|_{\ell_1})^* &= \sup\left\{\boldsymbol{x}^T\boldsymbol{y} + \|\boldsymbol{x}\|_{\ell_1} \;\middle|\; \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}}\right\} \\
&= \sup\left\{\boldsymbol{x}^T\boldsymbol{y} + \boldsymbol{x}^T\boldsymbol{z} \;\middle|\; \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}}, \|\boldsymbol{z}\|_{\ell_\infty} \le 1\right\} \\
&= \sup\left\{\boldsymbol{x}^T\boldsymbol{y} \;\middle|\; \boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}}\right\},
\end{aligned}
\tag{76}
$$

which clearly is unbounded if $\boldsymbol{x} \neq \boldsymbol{0}$. Therefore, we cannot formulate the dual to LP (67).

## A.3 Invexity

The KKT conditions are optimality conditions for non-linear optimization problems. Invexity is a generalization of convexity in the sense that the KKT conditions are necessary and sufficient for optimality [32]. By invexity we mean that the objective and constraint functions of the optimization problem are *Type 1 invex* functions.

**Definition 22.** *Consider the standard form of an optimization problem*

$$
\begin{aligned}
\text{minimize} \quad & f(\boldsymbol{y}) \\
\text{subject to} \quad & g(\boldsymbol{y}) \le \boldsymbol{0}, \\
& \boldsymbol{y} \in \mathcal{S},
\end{aligned}
\tag{77}
$$

*where $\mathcal{S} \subseteq \mathbb{R}^m$ is defined by $g(\boldsymbol{y}) \le \boldsymbol{0}$. Then $f$ and $g$ are called* Type 1 invex *functions at point $\boldsymbol{y}^* \in \mathcal{S}$ w.r.t. a common function $\eta(\boldsymbol{y}, \boldsymbol{y}^*) \in \mathbb{R}^m$, if for all $\boldsymbol{y} \in \mathcal{S}$,*

$$
\begin{aligned}
f(\boldsymbol{y}) - f(\boldsymbol{y}^*) &\ge \eta(\boldsymbol{y}, \boldsymbol{y}^*)^T \nabla f(\boldsymbol{y}^*), \\
-g(\boldsymbol{y}^*) &\ge \eta(\boldsymbol{y}, \boldsymbol{y}^*)^T \nabla g(\boldsymbol{y}^*)
\end{aligned}
\tag{78}
$$

*hold. It suffices to consider only the active constraints, i.e. the constraints, where equality holds $g(\boldsymbol{y}^*) = 0$. [32]*

Let $K$ be the scaling factor of the conjectured upper bound, i.e.

$$
K := \begin{cases} n\,, & \text{for odd } n\,, \\ n-1\,, & \text{for even } n\,. \end{cases}
\tag{79}
$$

In the case of LP (67) we have $\mathcal{S} = \left\{\boldsymbol{y} \in \mathbb{R}^{\binom{n}{2}} \;\middle|\; V^T\boldsymbol{y} \le \boldsymbol{1}\right\}$ and want to show that $\boldsymbol{y}^* = -\boldsymbol{1}/\lfloor n/2 \rfloor$ is a global optimum. Furthermore, we have

$$
\begin{aligned}
f(\boldsymbol{y}) &= -\|\boldsymbol{y}\|_{\ell_1}\,, \quad f(\boldsymbol{y}^*) = -K\,, \quad \nabla f(\boldsymbol{y}^*) = \boldsymbol{1}\,, \\
g(\boldsymbol{y}) &= V^T\boldsymbol{y} - \boldsymbol{1}\,, \quad g(\boldsymbol{y}^*) = 0\,, \quad \nabla g(\boldsymbol{y}^*) = V|_a\,,
\end{aligned}
\tag{80}
$$

where $V|_a$ are the columns of $V$ such that $(V|_a)^T\boldsymbol{y}^* = \boldsymbol{1}$, i.e. the active constraints. To show invexity we have to find a common $\eta(\boldsymbol{y}, \boldsymbol{y}^*) \in \mathbb{R}^m$ such that

$$
\begin{aligned}
\|\boldsymbol{y}\|_{\ell_1} &\le K - \eta^T(\boldsymbol{y}, \boldsymbol{y}^*)\boldsymbol{1} \ \text{ and} \\
(V^T|_a)\eta(\boldsymbol{y}, \boldsymbol{y}^*) &\le \boldsymbol{0}\,,
\end{aligned}
\tag{81}
$$

for all $\boldsymbol{y} \in \mathcal{S}$. It is quite challenging to find a $\eta(\boldsymbol{y}, \boldsymbol{y}^*) \in \mathbb{R}^m$ satisfying both inequalities. In particular, ansätze motivated from the geometry in small dimensions eventually fail for larger $n$.

## B Acronyms

**IQP** Instantaneous Quantum Polynomial time . . . . . . . . . . . . . . . . . . . . . . . . 2

**KKT** Karush–Kuhn–Tucker . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

**LP** linear program . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

**MIP** mixed integer program . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

**NISQ** noisy and intermediate scale quantum . . . . . . . . . . . . . . . . . . . . . . . . 2

## References


[1] X. Wang, A. Sørensen, and K. Mølmer, *Multibit gates for quantum computing*, Phys. Rev. Lett. **86**, 3907 (2001), arXiv:quant-ph/0012055.

[2] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, *14-qubit entanglement: Creation and coherence*, Phys. Rev. Lett. **106**, 130506 (2011), arXiv:1009.6126.

[3] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, *Superconducting qubits: Current state of play*, Annual Review of Condensed Matter Physics **11**, 369 (2020), arXiv:1905.13641.

[4] C. Figgatt, A. Ostrander, N. M. Linke, K. A. Landsman, D. Zhu, D. Maslov, and C. Monroe, *Parallel entangling operations on a universal ion-trap quantum computer*, Nature **572**, 368 (2019), arXiv:1810.11948.

[5] Y. Lu, S. Zhang, K. Zhang, W. Chen, Y. Shen, J. Zhang, J.-N. Zhang, and K. Kim, *Scalable global entangling gates on arbitrary ion qubits*, Nature **572**, 363 (2019), arXiv:1901.03508.

[6] P. Baßler, M. Zipper, C. Cedzich, M. Heinrich, P. H. Huber, M. Johanning, and M. Kliesch, *Synthesis of and compilation with time-optimal multi-qubit gates*, Quantum **7**, 984 (2023), arXiv:2206.06387.

[7] F. Barahona and A. R. Mahjoub, *On the cut polytope*, Mathematical Programming **36**, 157 (1986).

[8] M. R. Garey and D. S. Johnson, *Computers and intractability*, Vol. 29 (W. H. Freeman and Company, New York, 2002).

[9] M. J. Bremner, A. Montanaro, and D. J. Shepherd, *Average-case complexity versus approximate simulation of commuting quantum computations*, Phys. Rev. Lett. **117**, 080501 (2016), arXiv:1504.07999.

[10] J. Allcock, J. Bao, J. F. Doriguello, A. Luongo, and M. Santha, *Constant-depth circuits for Uniformly Controlled Gates and Boolean functions with application to quantum memory circuits*, arXiv:2308.08539 (2023).

[11] S. Bravyi, D. Maslov, and Y. Nam, *Constant-cost implementations of Clifford operations and multiply controlled gates using global interactions*, Phys. Rev. Lett. **129**, 230501 (2022), arXiv:2207.08691.

[12] S. Bravyi and D. Maslov, *Hadamard-free circuits expose the structure of the Clifford group*, IEEE Trans. Inf. Theory **67**, 4546 (2021), arXiv:2003.09412.

[13] D. Maslov and B. Zindorf, *Depth optimization of CZ, CNOT, and Clifford circuits*, IEEE Transactions on Quantum Engineering **3**, 1 (2022), arxiv:2201.05215.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2009).

[15] E. Rich, *The problem classes FP and FNP*, in *Automata, Computability and Complexity: Theory and Applications* (Pearson Education, 2007) pp. 510–511.

[16] M. Johanning, *Isospaced linear ion strings*, Appl. Phys. B **122**, 71 (2016).

[17] M. Laurent and S. Poljak, *On a positive semidefinite relaxation of the cut polytope*, Linear Algebra and its Applications **223-224**, 439 (1995).

[18] M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, 1st ed., Algorithms and Combinatorics (Springer Berlin Heidelberg, 2009).

[19] M. E.-Nagy, M. Laurent, and A. Varvitsiotis, *Complexity of the positive semidefinite matrix completion problem with a rank constraint*, Springer International Publishing , 105 (2013), arXiv:1203.6602.

[20] R. E. A. C. Paley, *On orthogonal matrices*, Journal of Mathematics and Physics **12**, 311 (1933).

[21] A. Hedayat and W. D. Wallis, *Hadamard matrices and their applications*, The Annals of Statistics **6**, 1184 (1978).

[22] H. Kharaghani and B. Tayfeh-Rezaie, *A Hadamard matrix of order 428*, Journal of Combinatorial Designs **13**, 435 (2005).

[23] D. Ž. Đoković, O. Golubitsky, and I. S. Kotsireas, *Some new orders of Hadamard and Skew-Hadamard matrices*, Journal of Combinatorial Designs **22**, 270 (2014), arXiv:1301.3671.

[24] J. Cohn, M. Motta, and R. M. Parrish, *Quantum filter diagonalization with compressed double-factorized Hamiltonians*, PRX Quantum **2**, 040352 (2021), arXiv:2104.08957.

[25] D. A. Spielman and S.-H. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, Journal of the ACM **51**, 385 (2004), arXiv:cs/0111050.

[26] S. Diamond and S. Boyd, *CVXPY: A Python-embedded modeling language for convex optimization*, J. Mach. Learn. Res. **17**, 1 (2016), arXiv:1603.00943.

[27] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, *A rewriting system for convex optimization problems*, J. Control Decis. **5**, 42 (2018), arXiv:1709.04494.

[28] Free Software Foundation, *GLPK (GNU Linear Programming Kit)* (2012), version: 0.4.6.

[29] A. T. Phillips and J. B. Rosen, *A parallel algorithm for constrained concave quadratic global minimization*, Mathematical Programming **42**, 421 (1988).

[30] M. Dür, R. Horst, and M. Locatelli, *Necessary and sufficient global optimality conditions for convex maximization revisited*, Journal of Mathematical Analysis and Applications **217**, 637 (1998).

[31] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: theory and algorithms, 3rd edition* (John wiley & sons, 2013).

[32] M. A. Hanson, *Invexity and the Kuhn–Tucker Theorem*, Journal of Mathematical Analysis and Applications **236**, 594 (1999).