General, Efficient, and Robust Hamiltonian Engineering

P. Baßler[1,2,*] M. Heinrich[1,3,†] and M. Kliesch[2,‡]

[1] *Heinrich Heine University Düsseldorf, Faculty of Mathematics and Natural Sciences, 40225 Düsseldorf, Germany*

[2] *Hamburg University of Technology, Institute for Quantum Inspired and Quantum Optimization, 21079 Hamburg, Germany*

[3] *University of Cologne, Institute for Theoretical Physics, 50937 Köln, Germany*

(Received 2 May 2025; accepted 15 October 2025; published 26 November 2025)

Implementing the time evolution under a desired target Hamiltonian is critical for various applications in quantum science. Due to the exponential increase of parameters in the system size and due to experimental imperfections, this task can be challenging in quantum many-body settings. We introduce an efficient and robust scheme to engineer arbitrary local many-body Hamiltonians. To this end, our scheme applies single-qubit π or $\pi/2$ pulses to an always-on system Hamiltonian, which we assume to be native to a given platform. These sequences are constructed by efficiently solving a linear program (LP) which minimizes the total evolution time. In this way, we can engineer target Hamiltonians that are only limited by the locality of the interactions in the system Hamiltonian. Based on average Hamiltonian theory and by using robust composite pulses, we make our schemes robust against errors, including finite-pulse-time errors and various control errors. To demonstrate the performance of our scheme, we provide numerical simulations. In particular, we solve the Hamiltonian-engineering problem on a laptop for arbitrary two-local Hamiltonians on a two-dimensional square lattice with 196 qubits in only 60 s. Moreover, we simulate the engineering of general Heisenberg Hamiltonians from Ising Hamiltonians with imperfect single-qubit pulses for smaller system sizes, and achieve a fidelity larger than 99.9%, which is orders of magnitude better than nonrobust implementations.

DOI: 10.1103/9yxv-tdqr

I. INTRODUCTION

Simulating a target Hamiltonian on a quantum system is a central problem in quantum computing, with applications in general gate-based and digital-analog quantum computing, as well as quantum simulations [1–6] and quantum chemistry [7–9]. It is commonly believed that simulating the dynamics of a quantum system is one of the most promising tasks for showing a practical advantage of quantum computers over classical computers [7,10,11], perhaps already on noisy intermediate-scale quantum (NISQ) devices [12,13]. Especially for the latter, it is essential to have an efficient and fast implementation of the target Hamiltonian due to short coherence times.

The idea of engineering a target Hamiltonian by interleaving the evolution under a fixed system Hamiltonian with single-qubit pulses originated in the nuclear magnetic resonance (NMR) community. However, the first approaches require additional single-qubit pulses to decouple interactions, rescale them, and couple them again, resulting in long pulse sequences [14–18]. Recently, there has been renewed interest in designing pulse sequences to change the effective dynamics governed by a given Hamiltonian. In particular, there has been impressive progress in the design of robust global pulses, identical pulses on each qubit, to change the global properties of a given Hamiltonian [19]. This approach has already been generalized to qudit systems [20,21] and implemented in experiments with ultracold atomic Rydberg gas and nitrogen-vacancy (NV) centers in diamond [22,23]. However, Hamiltonian-engineering schemes utilizing global pulses, such as Floquet Hamiltonian-engineering methods [22,24], cannot modify individual interaction terms. In our work, we design robust sequences of local pulses to change any interaction coupling in a given Hamiltonian, with similar robustness properties as for global pulses [19]. Moreover, due to the local pulses in our method we are able to

*Contact author: bassler@hhu.de
†Contact author: markus.heinrich@uni-koeln.de
‡Contact author: martin.kliesch@tuhh.de

TABLE I. A comparison of Hamiltonian-engineering methods for implementing a target Hamiltonian with r interactions on n qubits. We compare approaches similar to ours with respect to classical run-time and memory usage, the circuit depth (number of pulses), and the robustness. For general k-local Hamiltonians we have $r = \mathrm{O}(n^k)$. Thus, for a 2-local Hamiltonian we have $r = \mathrm{O}(n^2)$ and for a 2-local Hamiltonian on a one-dimensional (1D) spin chain we have $r = \mathrm{O}(n)$. Moreover, for a better comparison of methods we omit the overhead introduced by implementing noncommuting interactions, i.e., via product formulas, in the depth consideration.

Method	Target Hamiltonians	Run-time	Memory	Depth	Robust
Choi *et al.* [19]	Cannot modify individual interactions	$\mathrm{O}(1)$	$\mathrm{O}(1)$	$\mathrm{O}(1)$	Yes
DAQC [28]	2-local, arbitrary	$\mathrm{O}(4^n)$	$\mathrm{O}(4^n)$	$\mathrm{O}(n^2)$	No
Hayes *et al.* [29]	2-local, 1D spin chain	$\mathrm{O}(n^3)$[b]	$\mathrm{O}(n^2)$	$\mathrm{O}(n^2)$	No
EASE gate [32]	2-local, Ising type	$\mathrm{O}(n^3)$	$\mathrm{O}(n^2)$	$\mathrm{O}(n)$[a]	No
Votto *et al.* [18]	2-local, XY model	$\mathrm{O}(n^3)$	$\mathrm{O}(n^3)$	$\mathrm{O}(n^3)$	Yes
This work	Arbitrary (locality preserving)	$\mathrm{O}(r^3)$[b]	$\mathrm{O}(r^2)$	r	Yes

[a]Requires pulses of arbitrary angle.
[b]We have assumed a complexity of $\mathrm{O}(m^3)$ for solving a linear program with m variables [40].

implement the dynamics of a much larger family of possible target Hamiltonians. We believe that neutral-atom architectures based on different operating zones or in the weak-coupling regime are well suited for digital-analog methods as presented in our work [25–27].

So far, other approaches utilizing local pulses have certain limitations. They require NP-hard classical preprocessing to find the pulse sequence [28], rely on specific structures in the system Hamiltonian [29], or require an infinite single-qubit gate set which might be a problem for the fast control electronics in an experiment [30–32]. We provide a comparison of previous methods in Table I. These works are also limited to two-body interactions, and to the best of our knowledge, no general scheme has been developed for efficiently engineering many-body interactions individually. Since there has been an increasing effort in realizing the latter in experiments [4,33–36], such a general scheme would allow to simulate quantum chemistry Hamiltonians with genuine many-body interactions. Moreover, some recent Hamiltonian learning schemes rely on "reshaping" unknown many-body Hamiltonians to diagonal Hamiltonians which can be done efficiently and robust to errors with our proposed method [37–39].

In previous works, we have proposed a Hamiltonian engineering method for Ising Hamiltonians based on linear programming and applied it to multi-qubit gate synthesis and compiling problems [41,42]. A related linear program approach for commuting Ising and nearest-neighbor interactions with an efficient relaxation has been introduced in Refs. [43,44]. In this work, we generalize the linear-programming method to a significantly larger family of local Hamiltonians, make it substantially more efficient, and render it robust against dominant error sources. More concretely, our method allows us to efficiently engineer arbitrary target Hamiltonians, only limited by the locality of the system Hamiltonian. To this end, free evolutions under the system Hamiltonian are interleaved with π or $\pi/2$ pulses (i.e., Pauli or single-qubit Clifford gates). With the linear program (LP) formulation, we find an *exact* decomposition of the target Hamiltonian as a sum, where each term corresponds to single-qubit pulses and the corresponding evolution time under the system Hamiltonian. With this exact decomposition, one can implement the target evolution utilizing standard Hamiltonian-simulation methods such as Trotterization. A crucial feature of our method is that it minimizes the total evolution time, leading to a fast implementation of the target Hamiltonian. For larger systems, however, optimally solving the minimizing linear program is no longer efficiently possible. We introduce a relaxation that only scales with the number of interaction terms and not directly with the system size, and thus, provides an efficient method to engineer Hamiltonians. This relaxation still yields an exact decomposition of the target Hamiltonian, but the total evolution time may no longer be minimal. The latter can, however, be decreased by expanding the search space for the relaxed problem, providing a trade-off between the run-time of the classical preprocessing and the evolution time of the implementation.

Concretely, for any target Hamiltonian with r interaction terms and the same locality as the system Hamiltonian, our method efficiently finds an implementing single-qubit pulse sequence using a linear-in-r number of layers. To implement the target evolution with a product formula, we therefore also require $\mathrm{O}(r)$ single-qubit pulse layers, where the implicit constant only depends on the chosen product formula. Note that previous approaches require $\mathrm{O}(r^2)$ single-qubit pulse layers [14–18] or are classically hard to solve [28]. Moreover, all of these methods except Ref. [18] are not robust against any error source. For our method we observe that the total evolution time to implement the dynamics under the target Hamiltonian scales only sublinear with the number of qubits.

We also investigate the effects of dominant experimental error sources and introduce a general framework to cancel

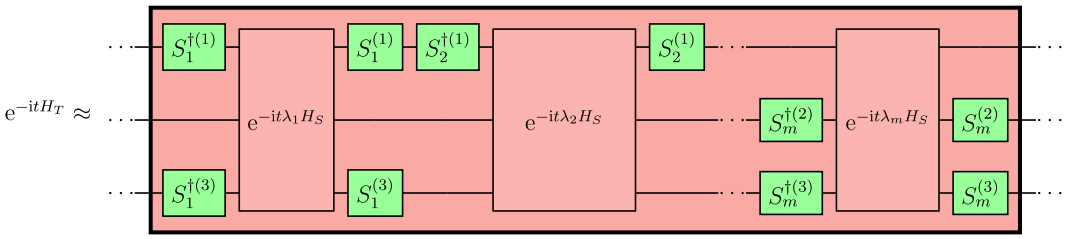


FIG. 1. We engineer the target Hamiltonian H_T by interleaving the natural dynamics of the quantum device, governed by the system Hamiltonian H_S, with layers of single-qubit π or $\pi/2$ pulses, $\boldsymbol{S}_i = S_i^{(1)} \otimes \ldots S_i^{(n)}$, as in Eq. (6). The large red box highlights that H_S is assumed to be always on. Our Hamiltonian-engineering results are straightforward to implement: apply single-qubit pulse layers $\boldsymbol{S}_{i-1}\boldsymbol{S}_i^\dagger$, let the system evolve freely under H_S for a duration $t\lambda_i$, and repeat.

them. To this end, we generalize a method by Votto *et al.* [18] based on average Hamiltonian theory (AHT) to mitigate these errors. As a result, we are able to make our methods robust against finite-pulse-time and single-qubit rotation-angle errors. We leverage the generality of our method to combine it with robust composite pulses, making it robust against many experimental error sources. This is especially beneficial in light of recent efforts which have demonstrated a remarkable single-qubit pulse fidelity of 10^{-6} utilizing fast robust composite pulses [45].

In summary, we propose a novel Hamiltonian-engineering method that is only limited by the locality of the native system interactions, in which the number of required pulses scales linear with the number of interactions and the total evolution time is minimized, and which is robust against common experimental imperfections. Therefore, we believe that our method finds many applications in medium-scale quantum simulation and gate-based quantum computation.

In Sec. II we provide an overview of our general, efficient, and robust methods. In Sec. III B we show the generality, efficiency, and robustness of our methods in numerical simulations of a realistic setting. In Sec. IV we introduce our general framework for engineering Hamiltonians using an LP. In Secs. IV A and IV B the efficient Hamiltonian-engineering methods by Pauli and Clifford conjugation are presented. Finally, in Sec. V we provide methods to investigate and mitigate the effect of errors on both the Pauli- and Clifford-conjugation methods.

II. OVERVIEW OF MAIN RESULTS

Before describing the technical details of our Hamiltonian-engineering method, we first provide an overview of our main contributions. We begin by introducing the general framework, then highlight its generality, efficiency, and robustness using the example of an Ising-type system Hamiltonian.

In the following, we denote the 2×2 identity and Pauli matrices as I, X, Y, and Z and use $\mathbb{1}_n \equiv \mathbb{1}$ for the n-qubit identity matrix. Tensor products of Pauli matrices are called *multi-qubit Pauli operators* or *Pauli strings*, the set of which is denoted as $\mathsf{P}^n = \{I, X, Y, Z\}^{\otimes n}$. We note that such a Pauli operator, e.g., $P = X \otimes Y \otimes X \otimes I$, can be both interpreted as a layer of simultaneous single-qubit π pulses and as an interaction term in a local Hamiltonian. This is because the Pauli operators form a basis for the space of Hermitian operators, and we can thus write any Hamiltonian as

$$H = \sum_{P \in \mathsf{P}^n} J_P P, \tag{1}$$

where P represents the interaction and $J_P \in \mathbb{R}$ is the corresponding interaction strength. We collect all J_P in a vector $\boldsymbol{J} \in \mathbb{R}^{4^n}$. We denote the *locality* of an interaction P, i.e., the qubit indices on which P acts nontrivially, as $\mathrm{supp}(P)$. For any constant k, we call the Hamiltonian in Eq. (1) *k-local* if $J_P = 0$ whenever $|\mathrm{supp}(P)| > k$.

Our goal is to simulate the time evolution under a desired target Hamiltonian H_T by using the native evolution of a system Hamiltonian H_S with layers of (parallel) single-qubit pulses $\boldsymbol{S}$. We express the system and target Hamiltonian as

$$H_S = \sum_{P \in \mathsf{P}^n} J_P P, \quad H_T = \sum_{P \in \mathsf{P}^n} A_P P. \tag{2}$$

We consider two types of pulse layers: π pulses (Pauli gates) and $\pi/2$ pulses (single-qubit Clifford gates). By conjugating the system Hamiltonian with the pulses $\boldsymbol{S}$, we change its dynamics as

$$\boldsymbol{S}^\dagger e^{-itH_S}\boldsymbol{S} = e^{-it\boldsymbol{S}^\dagger H_S \boldsymbol{S}}. \tag{3}$$

This transformation is also known as the toggling-frame transformation. Our objective is to find an exact decomposition of the target Hamiltonian such that

$$H_T = \sum_{i=1}^{r} \lambda_i \boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i, \tag{4}$$

where $\lambda_i > 0$ denotes the *relative evolution time* under the system Hamiltonian, $\boldsymbol{S}_i$ denotes the corresponding

single-qubit pulse layer, and r is the number of interaction terms that can be generated by the transformations. Below Eq. (LP), we explain why r coincides with the number of terms in the decomposition. The transformation $\boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i$ linearly transforms the interaction strengths $\{J_P\}$, and we capture this effect with a matrix $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$: each entry $\mathcal{W}(\boldsymbol{J})_{P,i}$ specifies the contribution of the interaction term P under the transformation with respect to the pulse layer i. Specifically, we have

$$\mathcal{W}(\boldsymbol{J})_{P,i} := \frac{1}{2^n} \operatorname{Tr}\left(P\left(\boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i\right)\right). \tag{5}$$

Thus, s is the number of considered single-qubit pulse layers.

Target Hamiltonians H_T that allow for a decomposition as in Eq. (4) need to have the same locality as the system Hamiltonian, i.e., we cannot generate interaction terms of higher locality from lower-locality ones (as we assume single-qubit pulses only). Moreover, the set of single-qubit pulses also restricts the possible target Hamiltonians; see the discussion in the following subsection.

If it exists, an exact decomposition as in Eq. (4) can be obtained by solving the following LP, which minimizes the total relative evolution time:

$$\begin{aligned} &\text{minimize} && \mathbf{1}^T\boldsymbol{\lambda} \\ &\text{subject to} && \mathcal{W}(\boldsymbol{J})\boldsymbol{\lambda} = \boldsymbol{A}, \boldsymbol{\lambda} \in \mathbb{R}^s_{\geq 0}, \end{aligned} \tag{LP}$$

where $\boldsymbol{A} \in \mathbb{R}^r$ represents the target interaction strengths and $\mathbf{1} = (1, 1, \ldots, 1)$ is the all-ones vector such that $\mathbf{1}^T\boldsymbol{\lambda} = \sum_i \lambda_i$. The number of considered pulse layers s equals the number of optimization variables λ_i, since each λ_i is associated with a specific $\boldsymbol{S}_i$ in Eq. (4). Equation (LP) is central for our approach to Hamiltonian engineering and its analysis plays an important role in this work. For this purpose, we require some basic notions and properties from the theory of linear programming [46] which we will briefly introduce in the following. A *feasible solution* $\boldsymbol{\lambda}$ is a vector which satisfies all constraints of Eq. (LP). An *optimal solution* is a feasible solution that also minimizes the objective function and is indicated by an asterisk as $\boldsymbol{\lambda}^*$. If Eq. (LP) has a feasible solution, then there also exists a r-sparse optimal solution $\boldsymbol{\lambda}^*$ [47], corresponding to a decomposition as in Eq. (4) with only r terms. Such a r-sparse optimal solution can be found using the simplex algorithm which, in practice, has a run-time that scales polynomial in the problem size $r \times s$ [40,48]. In Sec. IV, we derive conditions for the existence of a solution to this LP. In Secs. IV A 1 and IV B 1, we further demonstrate that choosing pulse layers $\boldsymbol{S}_i$ at random is an effective strategy for obtaining feasible LPs with relatively few variables.

The resulting dynamics are then implemented as a sequence of evolutions under H_S for time $t\lambda_i$ interleaved with single-qubit pulse layers $\boldsymbol{S}_i$:

$$\mathrm{e}^{-\mathrm{i}tH_T} \approx \boldsymbol{S}_1^\dagger \mathrm{e}^{-\mathrm{i}t\lambda_1 H_S} \boldsymbol{S}_1 \cdots \boldsymbol{S}_m^\dagger \mathrm{e}^{-\mathrm{i}t\lambda_m H_S} \boldsymbol{S}_m, \tag{6}$$

where m is the total number of applied evolution blocks. If the conjugated terms in Eq. (4) mutually commute, then the target evolution is implemented exactly, and we have $m = r$. Otherwise, the evolution can be approximated using, potentially higher-order, product formulas such as the Trotter formula; then we have $m > r$. We want to highlight that the only approximation of the target dynamics arises from the product formula, since our method provides an exact decomposition.

In the following discussion of this section we illustrate our method on a system with native Ising interactions of arbitrary coupling strengths

$$H_S = \sum_{i\neq j} J_{ij} Z_i Z_j, \tag{7}$$

with $J_{ij} \in \mathbb{R}$. Note that such a system Hamiltonian contains $n(n-1)/2$ distinct interaction terms.

A. Generality

A natural question to ask is which target Hamiltonians H_T can be generated by Eq. (4). In the following, we consider two different approaches: Pauli conjugation via π pulses (also known as refocusing pulses), and Clifford conjugation via $\pi/2$ pulses. We also assume that we have local control over individual qubits, thus we can apply arbitrary π or $\pi/2$ pulses to any subset of qubits. This level of control is necessary if we want to engineer arbitrary target coupling strengths. A lower level of control will restrict the possible target couplings but might still be sufficient in certain situations.

1. Pauli conjugation

This method allows the modification of individual coupling strengths in the system Hamiltonian. For an Ising system Hamiltonian (general Hamiltonians are discussed in Sec. IV A), we can engineer target Hamiltonians of the form

$$H_T = \sum_{i\neq j} A_{ij} Z_i Z_j, \tag{8}$$

with arbitrary interaction strength $A_{ij} \in \mathbb{R}$. This enables individual control over all $r = n/2(n-1)$ Ising interactions. Moreover, it is possible to modify interactions of unknown strength or even cancel them (see Sec. IV C). We illustrate this capability in Sec. III B 1, where pulse sequences are constructed to engineer an Ising Hamiltonian in the presence of additional, unknown three-body interaction terms. It is also possible to suppress certain long-range interactions in an Ising Hamiltonian to implement local multi-qubit gates.

2. Clifford conjugation

By using $\pi/2$ pulses, we can not only modify interaction strengths but also change interaction types, which we discuss in detail in Sec. IV B. In particular, the Ising system Hamiltonian can be engineered into arbitrary 2-local Hamiltonians of the form

$$H_T = \sum_{\substack{P\in\mathsf{P}^n \\ |\,\mathrm{supp}(P)|=2}} A_P P, \tag{9}$$

with arbitrary interaction strength $A_P \in \mathbb{R}$. This yields individual control over all $r = 3^2 n(n-1)/2$ distinct 2-local interactions. The prefactor 3^2 stems from the enlarged family of reachable target Hamiltonians. In general, for any k-body interaction in the system Hamiltonian, we can engineer 3^k distinct interaction terms with individually tunable strengths.

B. Efficiency

The LP described above involves r equality constraints, one per target interaction term, but the number of variables s required for an optimal solution scales exponentially with the number of qubits, i.e., $s = 4^n$ for π pulses or $s = 12^n$ for $\pi/2$ pulses. To address this, we propose an efficient relaxation that reduces the number of variables needed. In particular, it suffices to choose

$$s > 2r \tag{10}$$

variables to obtain a solution of the LP. Due to the relaxation the solution does not minimize the total evolution time. However, it provides a low-evolution-time solution that still yields an exact decomposition as in Eq. (4). This reduced LP allows a flexible trade-off between the classical run-time and the resulting quantum evolution time: increasing s reduces the evolution time at the cost of increased classical run-time. For more details on the efficient LP, see Secs. IV A 1 and IV B 1, and for a numerical investigation of the efficiency we refer to Sec. III A.

For Ising interactions, this leads to an exponential reduction in LP size, e.g., $s > n(n-1)$ for π pulses and $s > 3^2 n(n-1)$ for $\pi/2$ pulses.

C. Robustness

In practice, single-qubit pulses have some nonzero duration $t_p > 0$. Depending on the platform, the system Hamiltonian H_S remains active during that time. As a result, the implemented pulse and its inverse become

$$\boldsymbol{S}_{\mathrm{err}} = \mathrm{e}^{-\mathrm{i}t_p(H_S+H_c)} \quad \text{and} \quad \boldsymbol{S}'_{\mathrm{err}} = \mathrm{e}^{-\mathrm{i}t_p(H_S-H_c)}, \tag{11}$$

where H_c denotes the control-pulse Hamiltonian. Using AHT, we approximate the overall finite-pulse-time effect as

$$\boldsymbol{S}'_{\mathrm{err}}\mathrm{e}^{-\mathrm{i}tH_S}\boldsymbol{S}_{\mathrm{err}} \approx \mathrm{e}^{-\mathrm{i}t\boldsymbol{S}^\dagger H_S \boldsymbol{S}+H_{\mathrm{err}}}. \tag{12}$$

A key insight is that the error term H_{err} has the same locality as H_S in the first-order Magnus expansion and can, hence, be incorporated into the efficient LP framework. Therefore, the error H_{err} can be cancelled exactly (see Sec. V for more details).

Moreover, our framework supports the use of robust composite pulses in place of $\pi/2$ pulses. This enables mitigation of various control errors, including rotation-angle error (Rabi-frequency errors) and off-resonance errors [49,50]. We provide numerical simulations validating the robustness of our approach under realistic error models in Sec. III B.

III. NUMERICAL RESULTS

In this section, we assess the performance of our method by numerical simulations. We divide the analysis into two parts. First, we demonstrate the efficiency of our LP approach in terms of the total evolution time and classical run-time. This includes the scaling of the total evolution time for engineering general Hamiltonians (2-local, 3-local, and random many-body Hamiltonians), and the scaling of the classical run-time for 2-local Hamiltonians on a two-dimensional (2D) square lattice. Second, we simulate the time evolution under a target Hamiltonian, evaluating the robustness of our approach. These simulations include Ising interactions on a 2D lattice, unwanted three-body interactions, and control imperfections, as well as general Heisenberg Hamiltonians with imperfect pulse control.

Our robust LP methods can be replaced by a mixed-integer linear program (MILP), which we define in detail in Sec. V E, to reduce the number of required pulses at the cost of additional integer-valued constraints. All LPs and MILPs in our work are modeled using the PYTHON software package CVXPY [51,52] and, if not otherwise stated, solved with the MOSEK solver [53] on a workstation with 130 GB RAM and the AMD Ryzen Threadripper PRO 3975WX 32-Cores processor. Furthermore, we have made the code to reproduce all figures publicly available on GitHub [54].

A. Generality and efficiency of the LP approach

1. Scaling of the total evolution time

We solve the efficient Pauli and Clifford conjugation LPs, which utilize π and $\pi/2$ pulses and are defined in Secs. IV A and IV B, respectively. Both of these aim to reduce the total evolution time $\mathbf{1}^T\boldsymbol{\lambda} = \sum_i \lambda_i$. Increasing the number of considered single-qubit pulse layers s yields a larger search space for the LP, which typically finds a lower optimum and thus decreases the total evolution time.

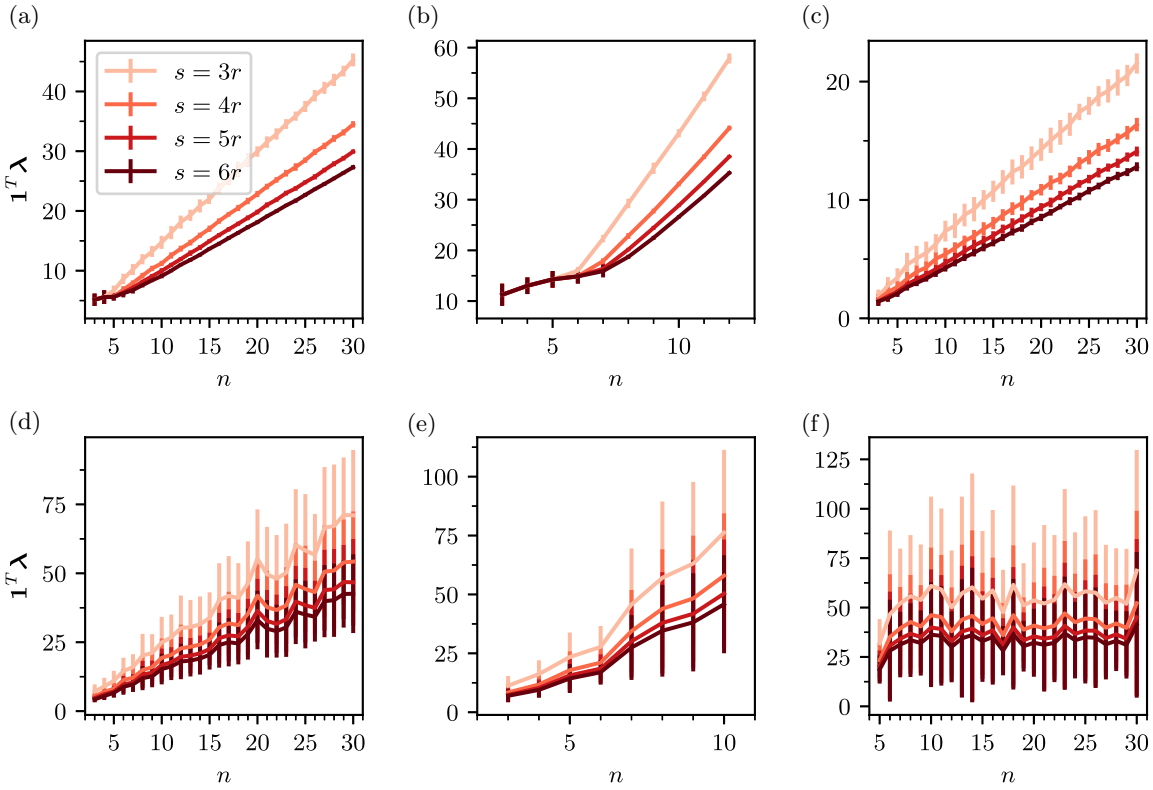


FIG. 2. The scaling of the total evolution time $\sum_i \lambda_i = \mathbf{1}^T\boldsymbol{\lambda}$ with system size n for randomly sampled target Hamiltonians. Each data point represents the average and the error bars represent the sample standard deviation over 50 uniform random samples of the target interaction strengths $\boldsymbol{A}$ from $[-1, 1]^r$. The numbers of variables in the linear program (LP) is varied from $s = 3r$ to $6r$ to demonstrate the reduction of the total evolution time when increasing the search space. (a)–(c) Results for the efficient Pauli conjugation LP with π pulses, using (a) 2-local, (b) 3-local, and (c) random many-body system Hamiltonians. (d)–(f) Results for the Clifford conjugation LP with $\pi/2$ pulses, using similar (d) 2-local and (e) 3-local system Hamiltonians, and (f) a randomly sampled 5-local Hamiltonian. Moreover, for (d) and (e) randomly selected interaction terms in the system Hamiltonian are set to zero.

We investigate how the total evolution time scales with the number of variables s and the system size n, considering various families of Hamiltonians with all-to-all connectivity, which are well suited for exploring the generality of our approach. Note that a k-local Hamiltonian with all-to-all connectivity contains $\sum_{i=2}^{k} 3^i \binom{n}{i}$ interaction terms. For instance, an Ising Hamiltonian includes $\binom{n}{2} = n(n-1)/2$ terms, while a general 2-local Hamiltonian has $3^2\binom{n}{2} = (9/2)n(n-1)$ terms.

In the Pauli-conjugation LP, the number of constraints r equals the number of nonzero terms in the system Hamiltonian. We solve instances of this LP with system Hamiltonians containing all possible 2-local interactions ($r = 3^2\binom{n}{2}$), all possible 3-local interactions ($r = 3^2\binom{n}{2} + 3^3\binom{n}{3}$), and randomly selected many-body interactions with $r = n^2$. The results are shown in Figs. 2(a)–2(c). In all cases, the total evolution time scales linearly with the system size, even with polynomial scaling in the number of interaction terms.

For the Clifford-conjugation LP, the number of constraints r, i.e., the number of interactions that can be generated, depends on both the number and locality of the interactions in the system Hamiltonian. We again consider 2-local and 3-local system Hamiltonians, as well as a random 5-local Hamiltonian. The random 5-local Hamiltonian is constructed by choosing ten random localities $\mathrm{supp}(P)$ for each $i = 2, \ldots, 5$ and including all 3^i interaction terms per locality, resulting in $r = 10\sum_{i=2}^{5} 3^i$ total terms which is independent of the system size. To add further variability, we randomly set some interaction strengths J_P to zero, ensuring that at least one term per locality remains nonzero. Nonzero interaction strengths are drawn independently from the uniform distribution, $J_P \overset{\text{IID}}{\sim} \mathrm{unif}([-1, +1])$ (where "IID" denotes "independent identically distributed"). Therefore, the number, index, and values of the nonzero interactions in the system Hamiltonians are random. In Figs. 2(d)–2(f), we show the scaling results for these systems. The evolution time again

scales roughly linearly for 2-local systems, slightly superlinearly for 3-local systems, and appears nearly constant for the random 5-local case with constantly many interactions. Due to the randomization, these simulations exhibit greater sample standard deviations compared to the Pauli-conjugation approach.

2. Evolution time versus classical run-time

To further assess practical efficiency, we evaluate the performance of our Pauli-conjugation method on 2-local Hamiltonians defined on a 2D square lattice. Such system Hamiltonians are in a broad sense motivated by experimental quantum platforms such as interacting superconducting qubits or cold atoms in optical lattices [55–57].

We consider the following system and target Hamiltonians:

$$H_S = J \sum_{\substack{P\in\mathsf{P}^n \\ \mathrm{supp}(P)\in E}} P, \quad H_T = \sum_{\substack{P\in\mathsf{P}^n \\ \mathrm{supp}(P)\in E}} A_P P, \tag{13}$$

where each P represents a two-body interaction term on a 2D square lattice with the set of edges E. For a $d\times d$ lattice the number of interactions is $r = 3^2|E| = 3^2 2d(d-1)$. The system interaction strength J is uniform across all interactions. This system Hamiltonian includes all possible interactions on the edges, i.e., interactions of the form Z_iY_j, X_iZ_j, or Y_iY_j with $(i,j)\in E$.

To emphasize scalability even on commodity hardware, we solve the efficient Pauli conjugation LP with the MOSEK solver [53] on a laptop with an Intel Core i7 Processor (8×1.8 GHz) and 16 GB RAM. As shown in Fig. 3, our method can engineer a target Hamiltonian H_T with arbitrary couplings $\boldsymbol{A}$ on 196 qubits in about 60 s of classical run-time using π pulses. Notably, the total evolution time scales sublinearly with the system size, allowing a fast implementation of the target Hamiltonian on platforms with restricted connectivity.

B. Numerical Hamiltonian simulations

To benchmark our methods, we perform numerical simulations that model the most relevant error sources occurring in practice. We consider a device with 2D lattice interactions, motivated by superconducting qubit platforms, and one with all-to-all connectivity modeling an ion trap. However, the presented methods are also applicable to other platforms such as cold atoms in the weak-coupling regime.

For simplicity, we describe the transformation of the system Hamiltonian H_S with respect to a single pulse layer. However, our simulations employ transformations of H_S containing multiple pulse layers, as detailed in the general discussion in Sec. V B. Let the ideal control Hamiltonian

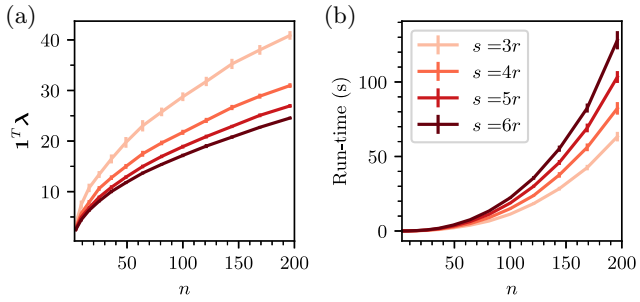


FIG. 3. A 2-local system Hamiltonian on a two-dimensional (2D) square lattice with $n = 4,\ldots,196$ qubits as in Eq. (13) and a constant interaction strength J. Each data point represents the average and the error bars represent the sample standard deviation over 50 uniform random samples of the target interaction strengths $\boldsymbol{A}$ from $[-1,1]^r$. (a) The evolution time over the number of qubits n. (b) The run-time for solving the Pauli conjugation LP over the number of qubits n.

for a single pulse layer on n qubits be given by

$$H_c = \frac{1}{t_p}\sum_{i=1}^{n} h_i, \tag{14}$$

where $h_i = (\pi/2)P_i$ for π pulses and $h_i = (\pi/4)P_i$ for $\pi/2$ pulses, with $P_i = I^{\otimes(i-1)}\otimes P\otimes I^{\otimes(n-i)}$ and $P\in\{I,X,Y,Z\}$. To model realistic errors, we introduce two sources of pulse imperfections by

$$H_{c,(\varepsilon,f)} = \frac{1}{t_p}\sum_{i=1}^{n}\left((1+\varepsilon_i)h_i + f_iZ_i\right), \tag{15}$$

where $\varepsilon_i \overset{\mathrm{IID}}{\sim} \mathrm{unif}([0,\varepsilon])$ is the *relative angle error* and $f_i \overset{\mathrm{IID}}{\sim} \mathrm{unif}([0,f])$ is the *off-resonance error* on the ith qubit. Both errors are sampled once and represent faulty and inhomogeneous control of the single-qubit pulses.

We also take the finite-pulse-time effects into account by modeling such single-qubit pulse layers:

$$\boldsymbol{S}_{\mathrm{err}} = \mathrm{e}^{-\mathrm{i}t_p(H_S+H_{c,(\varepsilon,f)})} \quad \text{and} \quad \boldsymbol{S}'_{\mathrm{err}} = \mathrm{e}^{-\mathrm{i}t_p(H_S-H_{c,(\varepsilon,f)})}, \tag{16}$$

which capture the joint evolution under the system Hamiltonian H_S and the imperfect control-pulse Hamiltonian $H_{c,(\varepsilon,f)}$. In the absence of H_S, these two operators are exact inverses of each other. We define the resulting erroneous *evolution block* as

$$U(t\lambda) = \boldsymbol{S}'_{\mathrm{err}}\mathrm{e}^{-\mathrm{i}t\lambda H_S}\boldsymbol{S}_{\mathrm{err}}, \tag{17}$$

which can be generalized to the expression given in Eq. (67) below.

We numerically simulate the time evolution given by the (robust) Pauli- and Clifford-conjugation methods by

explicitly computing products of matrix exponentials. More precisely, the decomposition of the target Hamiltonian is implemented using the first- or second-order Trotter formula,

$$U_{\mathrm{sim}} = \left(\overrightarrow{\prod_c} U\left(\frac{t\lambda_c}{n_{\mathrm{Tro}}}\right)\right)^{n_{\mathrm{Tro}}} \tag{18}$$

or

$$U_{\mathrm{sim}} = \left(\overrightarrow{\prod_c} U\left(\frac{t\lambda_c}{2n_{\mathrm{Tro}}}\right)\overrightarrow{\prod_c} U\left(\frac{t\lambda_c}{2n_{\mathrm{Tro}}}\right)\right)^{n_{\mathrm{Tro}}}, \tag{19}$$

where a single evolution block $U(t\lambda_c)$ consists of the time evolution under the system Hamiltonian conjugated by single-qubit pulses as in Eq. (67), and thus explicitly models finite-pulse-time errors. Throughout our work we use the convention $\overrightarrow{\prod}_{i=1}^{L} A_i := A_L \ldots A_1$ and $\overleftarrow{\prod}_{i=1}^{L} A_i := A_1 \ldots A_L$ to indicate the order of the products of noncommuting operators.

To capture the quality of the implementation, U_{sim} is compared to the target evolution $U_T = \mathrm{e}^{-\mathrm{i}tH_T}$. As a measure of quality, we use the *average gate infidelity*,

$$1 - F_{\mathrm{avg}}(U_{\mathrm{sim}}, U_T) = 1 - \frac{\mathrm{Tr}(U_T^\dagger U_{\mathrm{sim}}) + 1}{d+1}, \tag{20}$$

where d is the Hilbert-space dimension.

1. Simulation of a 2D lattice model

We consider a quantum platform with a native 2×3 lattice Hamiltonian with $n = 6$ qubits as an abstract model for interacting superconducting qubits:

$$H_S = J \sum_{ij} Z_i Z_j + \sum_{ijk} E_{ijk} X_i X_j X_k. \tag{21}$$

This Hamiltonian has $r = 17$ nonzero interaction terms. We assume that there are unwanted three-body interactions of unknown strength, where we consider all possible nearest-neighbor three-body interactions. Examples of the considered two- and three-body interactions are depicted in Fig. 4 as solid black lines and colored areas, respectively. The two-body coupling coefficients are constant, $J = 10^3$ Hz. The three-body coupling coefficients are uniformly sampled, $E_{ijk} \overset{\mathrm{IID}}{\sim} \mathrm{unif}([-10^2, 10^2] \cdot \mathrm{Hz})$, *after* the design of the pulse sequences and therefore considered as unknown. The finite π-pulse time is $t_p = 10^{-7}$ s. We want to implement the Ising Hamiltonian

$$H_T = \sum_{ij} A_{ij} Z_i Z_j, \tag{22}$$

with random coupling coefficients $A_{ij} \overset{\mathrm{IID}}{\sim} \mathrm{unif}([10^{-1}, 1] \cdot \mathrm{Hz})$. The pulse errors are modeled as in Eq. (15) with $\varepsilon = 10^{-1}$ and no off-resonance error.

In Fig. 4, we compare the NAIVE Pauli conjugation from Sec. IV A against the ROBUST PAULI conjugation from Sec. V D. For the sake of a clear presentation we have moved details to Appendix A 1. In the ANGLE ERROR and EXACT (TROTTER) data we apply the same sequences as for the NAIVE but with different errors (see the table in Fig. 4). To approximate the time evolution under H_T we use the first-order Trotter formula from Eq. (18) for all sequences. Increasing the number of Trotter cycles n_{Tro}

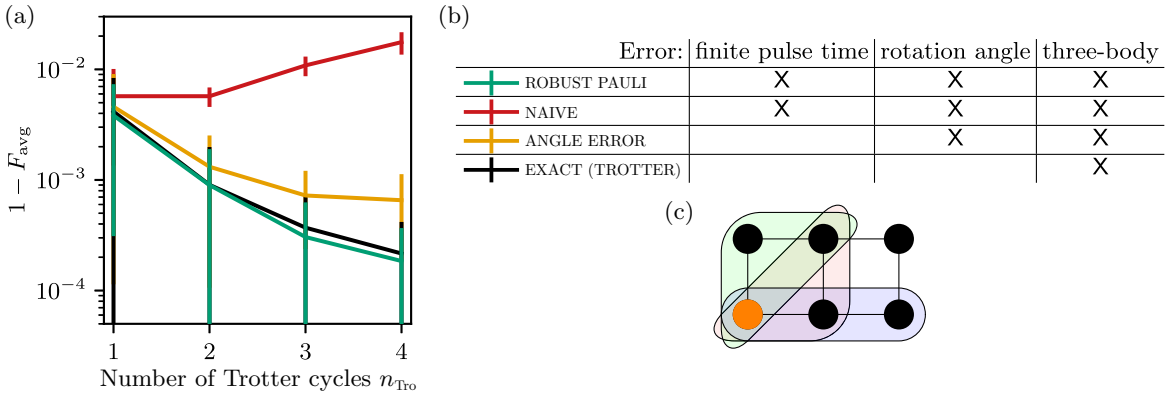


FIG. 4. (a) The sample mean and standard deviation of the average gate infidelities in Eq. (20) for implementing the time evolution of $\mathrm{e}^{-\mathrm{i}tH_T}$ with H_T from Eq. (22) and $t = 1$ s over the number of Trotter cycles n_{Tro}. The first-order Trotter formula from Eq. (18) is used to approximate the time evolution. The sample mean and standard deviation are calculated over 50 random samples of the target Hamiltonians H_T. Note that each of the n_{Tro} Trotter cycles contains κr evolution blocks $U(t\lambda_c)$, where $\kappa = 8$ and $s = 17$, with different rotation directions for the π pulses (see Sec. V D for more details). (b) A table indicating which error types are present for the different simulations. (c) An example 2×3 2D square lattice on $n = 6$ qubits with the two-body interactions (solid black lines) and the three-body interactions for the lower-left qubit (colored areas).

improves the accuracy of the Trotter approximation in the absence of other errors. However, an increased number of Trotter cycles n_{Tro} also yields an increased number of evolution blocks $U(t\lambda_c)$. From Fig. 4 it is clear that the finite-pulse-time errors and angle errors in the NAIVE sequences quickly accumulate even for a small number of Trotter cycles. However, the ROBUST PAULI sequences converge as quickly as the Trotter sequences without any errors.

2. Simulation of Heisenberg Hamiltonians with an ion-trap model

We consider an ion trap with ytterbium ions in an external magnetic field gradient [59]. The effective system Hamiltonian is

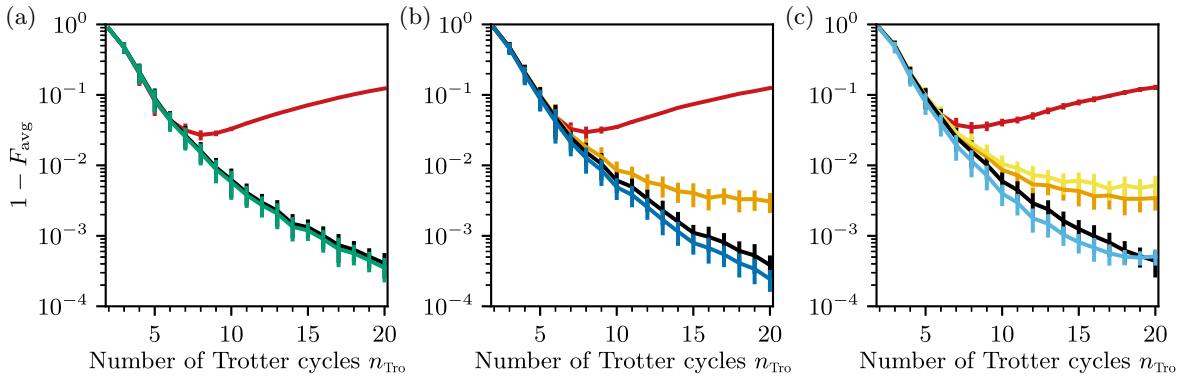
$$H_S = \sum_{i\neq j}^{n} J_{ij} Z_i Z_j , \tag{23}$$

where the coupling coefficients J_{ij} are calculated for a harmonic trapping potential affecting the equilibrium positions of the ions in the magnetic field gradient (cf. Ref. [41, App. A]). We consider all-to-all connectivity and thus the number of reachable interactions is $r = n/2(n-1)$ for the Pauli conjugation and $r = 3^2(n/2)(n-1)$ for the Clifford conjugation. The coupling coefficients J_{ij} are proportional to $(B_1/\omega)^2$, with a magnetic field gradient of $B_1 = 40T/m$ and a trap frequency of $\omega = 2\pi\,500$ kHz. The finite π-pulse time is $t_p = 2\ \mu$s, which is proportional to π/Ω with the Rabi frequency $\Omega = \omega/2$. Moreover, we consider the rotation-angle errors of strength $\varepsilon = 10^{-1}$ and off-resonance errors of strength $f = 10^{-1}$, which are modeled as in Eq. (15). As a target Hamiltonian, we consider the general Heisenberg Hamiltonian

$$H_T = \sum_{i\neq j}^{n} \left(A_{ij}^x X_i X_j + A_{ij}^y Y_i Y_j + A_{ij}^z Z_i Z_j \right), \tag{24}$$

with random coupling coefficients $A_{ij}^x, A_{ij}^y, A_{ij}^z \overset{\mathrm{IID}}{\sim} \mathrm{unif}([10^{-1}, 1]\cdot \mathrm{Hz})$ and all-to-all connectivity.

In Fig. 5 we compare the NAIVE Clifford conjugation from Sec. IV B to the ROBUST CLIFFORD, $\mathrm{CP}_{\mathrm{SCROFULOUS}}$ and $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ Clifford conjugations from Sec. V C. Again, we have moved details to Appendix A 2. The



(d)

Error:	finite pulse time	rotation angle	off-resonance
ROBUST CLIFFORD	X		
$\mathrm{CP}_{\mathrm{SCROFULOUS}}$	X	X	
$\mathrm{CP}_{\mathrm{SCORBUTUS}}$	X	X	X
NAIVE (LEFT)	X		
NAIVE (MIDDLE)	X	X	
NAIVE (RIGHT)	X	X	X
ANGLE ERROR		X	
OFF-RESONANCE ERROR			X
EXACT (TROTTER)			

FIG. 5. (a)–(c) The sample mean and standard deviation of the average gate infidelities in Eq. (20) for implementing the time evolution of e^{-itH_T} with H_T from Eq. (24) and $t = 1$ s over the number of Trotter cycles n_{Tro}. The second-order Trotter formula from Eq. (19) is used to approximate the time evolution. The sample mean and standard deviation are calculated over 50 random samples of the Heisenberg Hamiltonians H_T on $n = 8$ qubits. (a) The Clifford conjugation robust against finite-pulse-time errors (dark green), compared to the nonrobust Clifford conjugation (red). (b) The Clifford conjugation in combination with the SCROFULOUS ("Short Composite ROtation For Undoing Length Over and Under Shoot") pulse sequence [58] robust against finite-pulse-time errors and rotation-angle errors (dark blue), compared to the nonrobust Clifford conjugation (red). (c) The Clifford conjugation in combination with the SCORBUTUS ("Short COmposite Rotation BUffering Two Undesirables with Switchback") pulse sequence [50] robust against finite-pulse-time errors, rotation-angle errors, and off-resonance errors (light blue), compared to the nonrobust Clifford conjugation (red). (d) A table indicating which error types are present for the different simulations.

$\mathrm{CP}_{\mathrm{SCROFULOUS}}$ and $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ conjugations implement the robust composite pulses SCROFULOUS [58] and SCORBUTUS [50], respectively. These robust composite pulses are designed to compensate for angle errors (SCROFULOUS) or both angle errors and off-resonance errors (SCORBUTUS). For the ANGLE ERROR, OFF-RESONANCE ERROR and EXACT (TROTTER) data we apply the same sequences as for the NAIVE data but with different errors (see the table in Fig. 5). To approximate the time evolution under H_T we use the second-order Trotter formula from Eq. (19) for all sequences. As for the Pauli-conjugation method in Sec. III B 1, we can observe that the finite-pulse-time errors and angle errors in the NAIVE sequences quickly accumulate even for a moderate number of Trotter cycles. However, the ROBUST CLIFFORD, $\mathrm{CP}_{\mathrm{SCROFULOUS}}$ and $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ sequences converge as quickly as the Trotter sequences without any errors, and seem to be even more accurate at moderate numbers of Trotter cycles.

Note that the rotation-angle-error robustness in the $\mathrm{CP}_{\mathrm{SCROFULOUS}}$ Clifford conjugation is stronger than in the ROBUST PAULI conjugation in Sec. III B 1, since each $\pi/2$ pulse (replaced by robust composite pulses) is robust against different angle errors, whereas in the ROBUST PAULI conjugation the angle errors are assumed to be constant over several implementations of $U(t\lambda_c)$.

C. Summary of numerical results

This section presents a comprehensive numerical demonstration of the performance of our Hamiltonian engineering framework. We show that our methods achieve linear or sublinear scaling in evolution time with the system size, particularly for systems with constrained connectivity. We construct pulse sequences for engineering a 2D lattice Hamiltonian with 196 qubits in about 60 s on a laptop, highlighting the efficiency of our approach. Moreover, we show how our methods can be tailored for robustness against a variety of realistic control errors, including finite pulse durations, angle errors, and off-resonance effects. In summary, these numerical results underpin our claim of a general, efficient, and robust method to engineer Hamiltonians.

IV. HAMILTONIAN ENGINEERING BY LINEAR PROGRAMMING

In Sec. II we have already introduced our LP approach for Hamiltonian engineering. Now, we provide the technical details in full generality.

We begin by introducing a useful representation of Pauli operators as binary strings. $\mathbb{F}_2$ is the finite field over two elements, i.e., the set $\{0,1\}$, where multiplication and addition are performed modulo 2. We index each n-qubit Pauli operator $P_{\boldsymbol{a}} \in \mathsf{P}^n$ by a binary vector $\boldsymbol{a} = (\boldsymbol{a}_x, \boldsymbol{a}_z) \in \mathbb{F}_2^{2n}$ such that

$$P_{\boldsymbol{a}} = P_{(\boldsymbol{a}_x,\boldsymbol{a}_z)} = \mathrm{i}^{\boldsymbol{a}_x\cdot\boldsymbol{a}_z} X(\boldsymbol{a}_x) Z(\boldsymbol{a}_z). \tag{25}$$

Here, $X(\boldsymbol{x}) := X^{x_1} \otimes \cdots \otimes X^{x_n}$ and $Z(\boldsymbol{z}) := Z^{z_1} \otimes \cdots \otimes Z^{z_n}$, where X and Z are the single-qubit Pauli matrices. Sometimes, we also call $P_{\boldsymbol{a}}$ a *Pauli string*. For example, $P_{\boldsymbol{a}} = X \otimes Y \otimes X \otimes I$ has the binary representation $\boldsymbol{a} = ((1,1,1,0),(0,1,0,0)) \equiv (1,1,1,0,0,1,0,0)$. The Pauli strings satisfy the commutator relation

$$P_{\boldsymbol{a}} P_{\boldsymbol{b}} = (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle} P_{\boldsymbol{b}} P_{\boldsymbol{a}}, \tag{26}$$

for any $\boldsymbol{a},\boldsymbol{b} \in \mathbb{F}_2^{2n}$, with the binary symplectic form on $\mathbb{F}_2^{2n}$ defined by $\langle \boldsymbol{a},\boldsymbol{b}\rangle := \boldsymbol{a}_x \cdot \boldsymbol{b}_z + \boldsymbol{a}_z \cdot \boldsymbol{b}_x$, where the sum is taken in $\mathbb{F}_2$, i.e., modulo 2. With this notation, the locality of a Pauli operator $P_{\boldsymbol{a}}$ is defined as $\mathrm{supp}(P_{\boldsymbol{a}}) := \mathrm{supp}(\boldsymbol{a}) := \{i \in [n] | a_{x_i} \neq 0 \text{ or } a_{z_i} \neq 0\}$.

We write H_S and H_T in their Pauli decompositions, utilizing the binary notation

$$H_S = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} J_{\boldsymbol{a}} P_{\boldsymbol{a}}, \quad H_T = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} A_{\boldsymbol{a}} P_{\boldsymbol{a}}, \tag{27}$$

with $J_{\boldsymbol{a}}, A_{\boldsymbol{a}} \in \mathbb{R}$. We exclude the term $P_{\boldsymbol{0}} = P_{(0,\ldots,0)} = I^{\otimes n}$, which leads to a global and thus unobservable phase. Our goal is to simulate the time evolution under H_T by the one under H_S interleaved with layers of single-qubit pulses. Recall that we utilize the linear transformation $\mathrm{e}^{-\mathrm{i}tU^\dagger H U} = U^\dagger \mathrm{e}^{-\mathrm{i}tH} U$, which holds for any Hamiltonian H and unitary U. Hence, we seek a decomposition of the form in Eq. (4), with terms $\lambda_i \boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i$, where $\lambda_i > 0$ and $\boldsymbol{S}_i = S_i^{(1)} \otimes \cdots \otimes S_i^{(n)}$ representing a layer of single-qubit pulses, which we later set to be either π or $\pi/2$ pulses (i.e., Pauli or Clifford gates). To this end, we consider the effect of $H_S \mapsto \boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i$, similar to Eq. (5), on the Pauli coefficients $J_{\boldsymbol{a}}$ which can be captured by a column vector $\mathcal{W}(\boldsymbol{J})_i \in \mathbb{R}^r$ with elements

$$\mathcal{W}(\boldsymbol{J})_{\boldsymbol{a},i} := \frac{1}{2^n} \mathrm{Tr}\left(P_{\boldsymbol{a}} \left(\boldsymbol{S}_i^\dagger H_S \boldsymbol{S}_i\right)\right). \tag{28}$$

Here, r is the number of interaction terms $\{P_{\boldsymbol{a}}\}$ that we can generate from the ones in H_S by any pulse layer $\boldsymbol{S}_i$. Let s be the number of possible pulse layers on n qubits, and define the matrix $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$ by taking all vectors $\mathcal{W}(\boldsymbol{J})_i$ as columns. Then, a decomposition as in Eq. (4) can be found by solving Eq. (LP) minimizing the total *relative evolution time*. Conjugation of H_S with arbitrary single-qubit pulses yields a r which is only limited by the locality of the interactions in H_S. Below, we restrict ourselves to $\pi/2$ pulses achieving the same flexibility.

The first main contribution of our work is a hierarchy of relaxations of Eq. (LP) with drastically reduced size that still allows for an exact decomposition as in Eq. (4). For

this, we need general sufficient conditions on the matrix $\mathcal{W}(\boldsymbol{J})$ such that Eq. (LP) has a feasible solution for any $\boldsymbol{A} \in \mathbb{R}^r$.

Definition 1. We say that a matrix $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$ is *feasible* if for each $\boldsymbol{A} \in \mathbb{R}^r$ there exists a $\boldsymbol{\lambda} \in \mathbb{R}^s_{\geq 0}$ such that $\mathcal{W}(\boldsymbol{J})\boldsymbol{\lambda} = \boldsymbol{A}$.

This definition captures the constraints in Eq. (LP). There is a simple sufficient condition for the feasibility of the matrix $\mathcal{W}(\boldsymbol{J})$, as follows.

Proposition 1. Let $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$ such that $\ker(\mathcal{W}(\boldsymbol{J})^T) = \{\boldsymbol{0}\}$. If there exists $\boldsymbol{x} \in \mathbb{R}^s$ such that $\mathcal{W}(\boldsymbol{J})\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{x} > \boldsymbol{0}$, then for any $\boldsymbol{A} \in \mathbb{R}^r$ there exists $\boldsymbol{x}' \in \mathbb{R}^s$ such that $\mathcal{W}(\boldsymbol{J})\boldsymbol{x}' = \boldsymbol{A}$ and $\boldsymbol{x}' \geq \boldsymbol{0}$.

This proposition follows directly from two well-known results in convex analysis.

Lemma 1 (Farkas [60]). Let $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$ and $\boldsymbol{A} \in \mathbb{R}^r$. Then, exactly one of the following assertions is true:

(1) $\exists \boldsymbol{x} \in \mathbb{R}^s$ such that $\mathcal{W}(\boldsymbol{J})\boldsymbol{x} = \boldsymbol{A}$ and $\boldsymbol{x} \geq \boldsymbol{0}$.
(2) $\exists \boldsymbol{y} \in \mathbb{R}^r$ such that $\mathcal{W}(\boldsymbol{J})^T\boldsymbol{y} \geq \boldsymbol{0}$ and $\boldsymbol{A}^T\boldsymbol{y} < \boldsymbol{0}$.

Lemma 2 (Stiemke [61]). Let $\mathcal{W}(\boldsymbol{J}) \in \mathbb{R}^{r\times s}$. Then, exactly one of the following assertions is true:

(1) $\exists \boldsymbol{x} \in \mathbb{R}^s$ such that $\mathcal{W}(\boldsymbol{J})\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{x} > \boldsymbol{0}$.
(2) $\exists \boldsymbol{y} \in \mathbb{R}^r$ such that $\mathcal{W}(\boldsymbol{J})^T\boldsymbol{y} \gneq \boldsymbol{0}$.

Proof of proposition 1. We show that the second assertion of the Farkas lemma is not possible. By Lemma 2, there does not exist a $\boldsymbol{y}$ such that $\mathcal{W}(\boldsymbol{J})^T\boldsymbol{y} \gneq \boldsymbol{0}$. From $\ker(\mathcal{W}(\boldsymbol{J})^T) = \{\boldsymbol{0}\}$ it follows that if $\mathcal{W}(\boldsymbol{J})^T\boldsymbol{y} = \boldsymbol{0}$, then $\boldsymbol{y} = \boldsymbol{0}$, which contradicts $\boldsymbol{A}^T\boldsymbol{y} < \boldsymbol{0}$ in the second assertion of Lemma 1. Finally, the case $\mathcal{W}(\boldsymbol{J})^T\boldsymbol{y} < \boldsymbol{0}$ directly contradicts the second assertion of Lemma 1. ■

To provide a geometric interpretation of Proposition 1 we first define the *convex hull of the column vectors* of $\mathcal{W}(\boldsymbol{J})$,

$$\operatorname{conv}(\mathcal{W}(\boldsymbol{J})) := \left\{\boldsymbol{u} \in \mathbb{R}^r \middle| \mathcal{W}(\boldsymbol{J})\boldsymbol{x} = \boldsymbol{u}, \boldsymbol{x} \geq \boldsymbol{0}, \sum_i x_i = 1\right\}, \tag{29}$$

and the *interior of a polytope P*,

$$\operatorname{int}(P) := \left\{\boldsymbol{u} \in P \middle| \exists \varepsilon > 0 \text{ such that } \|\boldsymbol{u} - \boldsymbol{x}\| < \varepsilon \ \ \forall \boldsymbol{x} \in \mathbb{R}^r \Rightarrow \boldsymbol{x} \in P\right\}. \tag{30}$$

Equation (LP) is feasible if the convex hull of the column vectors of $\mathcal{W}(\boldsymbol{J})$ has an interior that contains the origin, $\boldsymbol{0} \in \operatorname{int}(\operatorname{conv}(\mathcal{W}(\boldsymbol{J})))$. The conditions of Proposition 1 can be efficiently verified. If the LP

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{1}^T\boldsymbol{x} \\ \text{subject to} \quad & \mathcal{W}(\boldsymbol{J})\boldsymbol{x} = \boldsymbol{0} \\ & \boldsymbol{x} \geq \boldsymbol{1} \end{aligned} \tag{31}$$

has a feasible solution and $\mathcal{W}(\boldsymbol{J})$ has full rank, then $\mathcal{W}(\boldsymbol{J})$ is feasible. Note that Proposition 1 applies to general matrices and will be useful for the efficient relaxations.

A. Pauli conjugation

Conjugation of a system Hamiltonian H_S with a Pauli string $P_{\boldsymbol{b}}$, i.e., a π-pulse layer, leads to

$$P_{\boldsymbol{b}}^\dagger H_S P_{\boldsymbol{b}} = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle} J_{\boldsymbol{a}} P_{\boldsymbol{a}}, \tag{32}$$

which follows from the commutation relations in Eq. (26). We seek a decomposition of H_T of the form

$$\begin{aligned} H_T &= \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} A_{\boldsymbol{a}} P_{\boldsymbol{a}} \\ &= \sum_{\boldsymbol{b}\in\mathbb{F}_2^{2n}} \lambda_{\boldsymbol{b}} P_{\boldsymbol{b}}^\dagger H_S P_{\boldsymbol{b}} \\ &= \sum_{\boldsymbol{a}\neq\boldsymbol{0},\boldsymbol{b}} \lambda_{\boldsymbol{b}} (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle} J_{\boldsymbol{a}} P_{\boldsymbol{a}}, \quad (\lambda_{\boldsymbol{b}} \geq 0). \end{aligned} \tag{33}$$

The $4^n \times 4^n$ matrix with entries $W_{\boldsymbol{ab}} = (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle}$ is called the *Walsh-Hadamard matrix*. Here, we only need a submatrix $W^{(r\times s)} \in \{-1, 1\}^{r\times s}$ defined by choosing r row indices $\boldsymbol{a} \in \mathbb{F}_2^{2n}$ and s column indices $\boldsymbol{b} \in \mathbb{F}_2^{2n}$. We call $W^{(r\times s)}$ a *partial Walsh-Hadamard matrix*. By Eq. (33), it is clear that only the interaction terms with $J_{\boldsymbol{a}} \neq 0$ contribute to the target Hamiltonian H_T. Therefore, we define $\operatorname{nz}(\boldsymbol{J}) := \left\{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\} \middle| J_{\boldsymbol{a}} \neq 0\right\}$, and require that the target interaction strengths satisfy

$$\operatorname{nz}(\boldsymbol{A}) \subseteq \operatorname{nz}(\boldsymbol{J}). \tag{34}$$

This requirement can be eliminated with single-qubit Clifford conjugation, which we will discuss in detail in Sec. IV B. In addition to the restriction $\boldsymbol{a} \in \operatorname{nz}(\boldsymbol{J})$ given by the system Hamiltonian, we may also restrict the set of Pauli strings (π-pulse layers) $\boldsymbol{b} \in \mathcal{F} \subseteq \mathbb{F}_2^{2n}$, as long as there is still a solution to Eq. (33). Comparing the target interaction strengths in Eq. (33), we can write this as

$$A_{\boldsymbol{a}} = J_{\boldsymbol{a}} \sum_{\boldsymbol{b}\in\mathcal{F}} W^{(r\times s)}_{\boldsymbol{ab}} \lambda_{\boldsymbol{b}} \quad \forall \boldsymbol{a} \in \operatorname{nz}(\boldsymbol{J}), \tag{35}$$

with the number of nonzero system interaction strengths $r := |\operatorname{nz}(\boldsymbol{J})|$ and the number of considered Pauli strings

$s = |\mathcal{F}|$. Hence, the constraint in Eq. (LP) becomes $\boldsymbol{A} = \boldsymbol{J} \odot (W^{(r\times s)}\boldsymbol{\lambda})$, where $\odot$ denotes element-wise multiplication and $\boldsymbol{A}, \boldsymbol{J} \in \mathbb{R}^r$ are restricted to $\boldsymbol{a} \in \mathrm{nz}(\boldsymbol{J})$. For the following analysis, it is convenient to set $\boldsymbol{M} := \boldsymbol{A} \oslash \boldsymbol{J}$ using the element-wise division $(\boldsymbol{A} \oslash \boldsymbol{J})_{\boldsymbol{a}} := A_{\boldsymbol{a}}/J_{\boldsymbol{a}}$ for all $\boldsymbol{a} \in \mathrm{nz}(\boldsymbol{J})$. This results in the following LP:

$$\begin{aligned} &\text{minimize} && \mathbf{1}^T\boldsymbol{\lambda} \\ &\text{subject to} && W^{(r\times s)}\boldsymbol{\lambda} = \boldsymbol{M},\ \boldsymbol{\lambda} \in \mathbb{R}^s_{\geq 0}. \end{aligned} \tag{PauliLP}$$

The run-time of an LP scales with its size, i.e., the number of constraints r and the number of variables s. We discuss the existence of a solution $\boldsymbol{\lambda} \in \mathbb{R}^{4^n}_{\geq 0}$ and bounds on the total relative evolution time $\mathbf{1}^T\boldsymbol{\lambda}$ in detail in Appendix B. Here, $r = |\mathrm{nz}(\boldsymbol{J})| \leq 4^n - 1$ is fixed by the system Hamiltonian and *a priori* $s = 4^n$; thus solving Eq. (PauliLP) already becomes computationally intensive at moderate system size. To overcome this, we propose a simple and efficient relaxation of Eq. (PauliLP) in Sec. IV A 1: for any $r = |\mathrm{nz}(\boldsymbol{J})|$, we sample $s \geq 2r$ Pauli strings at random, leading to a feasible LP with high probability. As we have $r = \mathsf{poly}(n)$ for local Hamiltonians, this indeed yields an efficient relaxation in the system size n. Note that this relaxation still leads to an exact decomposition; however, the relative evolution time $\mathbf{1}^T\boldsymbol{\lambda}$ may no longer be minimal.

1. Efficient relaxation

We provide an efficient method to construct a feasible matrix $W^{(r\times s)}$ with $r = |\mathrm{nz}(\boldsymbol{J})|$ rows and $s \geq 2r$ columns. The construction is rather simple and consists of sampling s vectors $\boldsymbol{b} \in \mathbb{F}_2^{2n}$ uniformly at random and taking the corresponding partial Walsh-Hadamard matrix with entries $W^{(r\times s)}_{\boldsymbol{ab}} = (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle}$, where $\boldsymbol{a} \in \mathrm{nz}(\boldsymbol{J})$. Thus, engineering a Hamiltonian with r nonzero interaction terms leads to Eq. (PauliLP), the number of constraints and variables in which both scale linearly with r.

To ensure feasibility of the subsampled matrix, we invoke Proposition 1. Thus, we have to check that $\mathrm{conv}(W^{(r\times s)})$ has a nonempty interior and $\mathbf{0} \in \mathrm{conv}(W^{(r\times s)})$. First, we state general results for IID copies $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s$ of a random vector $\boldsymbol{x}$. Then, we relate the results to the partial Walsh-Hadamard matrix $W^{(r\times s)}$. Here, we assume that s is large enough and that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s$ are in general positions such that $\mathrm{conv}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$ always has a nonempty interior. Thus, we focus on the condition $\mathbf{0} \in \mathrm{conv}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$.

Definition 2. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s$ be IID copies of an arbitrary random vector $\boldsymbol{x}$ in $\mathbb{R}^r$ and define

$$p_{s,\boldsymbol{x}} := \mathbb{P}(\mathbf{0} \in \mathrm{conv}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)). \tag{36}$$

Intuitively, one would expect that the probability should increase quickly with the number of samples s, and that taking s of the order of r should make $p_{s,\boldsymbol{x}}$ reasonably large.

Lemma 3 ([62], Proposition 4). Let $\boldsymbol{x} \in \mathbb{R}^r$ be an arbitrary random vector with $\mathbb{E}[\boldsymbol{x}] = 0$ and $\mathbb{P}(\boldsymbol{x} \neq 0) > 0$. Then, we have

$$0 < p_{r+1,\boldsymbol{x}} < p_{r+2,\boldsymbol{x}} < \cdots < p_{r+l,\boldsymbol{x}} \to 1 \quad \text{for } l \to \infty, \tag{37}$$

and $p_{s,\boldsymbol{x}} = 0$ if $s \leq r$.

Indeed, letting $\boldsymbol{w}$ be the r-dimensional random vector drawn uniformly from the columns of the partial Walsh-Hadamard matrix $W^{(r\times 4^n)}$, we can readily verify $\mathbb{E}[\boldsymbol{w}] = 0$ and thus Lemma 3 applies. However, the question of how large s should be taken still remains. In the convex-geometry literature, we find an elegant solution, at least for spherically symmetric random vectors, in the form of *Wendel's theorem* [63]: if the random vector $\boldsymbol{x}$ has a spherically symmetric distribution around $\mathbf{0}$, then the probability in Eq. (36) is given by

$$p_{s,\boldsymbol{x}} = 1 - \frac{1}{2^{s-1}} \sum_{k=0}^{r-1} \binom{s-1}{k}. \tag{38}$$

This distribution shows a sharp transition from ≈ 0 to ≈ 1 at $s = 2r$. However, Wendel's theorem does not apply to the random columns $\boldsymbol{w}$ of $W^{(r\times 4^n)}$ because of the lack of spherical symmetry (see Appendix C for details). Instead, the result in Eq. (38) of Wendel's theorem provides an upper bound on the probabilities $p_{s,\boldsymbol{w}}$ [62,64], and adequate lower bounds prove to be tricky to derive.

Nevertheless, we numerically observe that the random columns $\boldsymbol{w}$ of $W^{(r\times 4^n)}$ follow the behavior that we would expect from Wendel's theorem: we observe a sharp transition of $p_{s,\boldsymbol{w}}$ at the same position as for spherical symmetric distributions, approximating the upper bound in Eq. (38). In this sense, $p_{s,\boldsymbol{w}}$ is optimal, maximizing the success probability for finding a feasible $W^{(r\times s)}$.

Numerical observation 1 (Fig. 6). Let $W^{(r\times 4^n)}$ be a partial Walsh-Hadamard matrix with $r = |\mathrm{nz}(\boldsymbol{J})|$, and let $\boldsymbol{w}$ be a r-dimensional random vector drawn uniformly from $\mathrm{col}(W^{(r\times 4^n)})$. Then, we numerically observe that Wendel's statement in Eq. (38) approximately holds, i.e., the random submatrix $W^{(r\times s)}$ is feasible with high probability provided that we take $s \geq 2r$ samples of $\boldsymbol{w}$.

Let $W^{(r\times s)}$ be the partial Walsh-Hadamard matrix obtained from $W^{(r\times 4^n)}$ by randomly sampling s columns $\boldsymbol{b} \overset{\text{IID}}{\sim} \mathrm{unif}(\mathbb{F}_2^{2n})$ with $s \geq 2r$. As $W^{(r\times s)}$ is feasible with high probability, it can be used in Eq. (PauliLP) to engineer any target Hamiltonian with a Pauli decomposition compatible with the system Hamiltonian, i.e., with $\mathrm{nz}(\boldsymbol{A}) \subseteq \mathrm{nz}(\boldsymbol{J})$. One $W^{(r\times s)}$ can be reused for different Hamiltonians with a Pauli decomposition with r terms, since the construction of

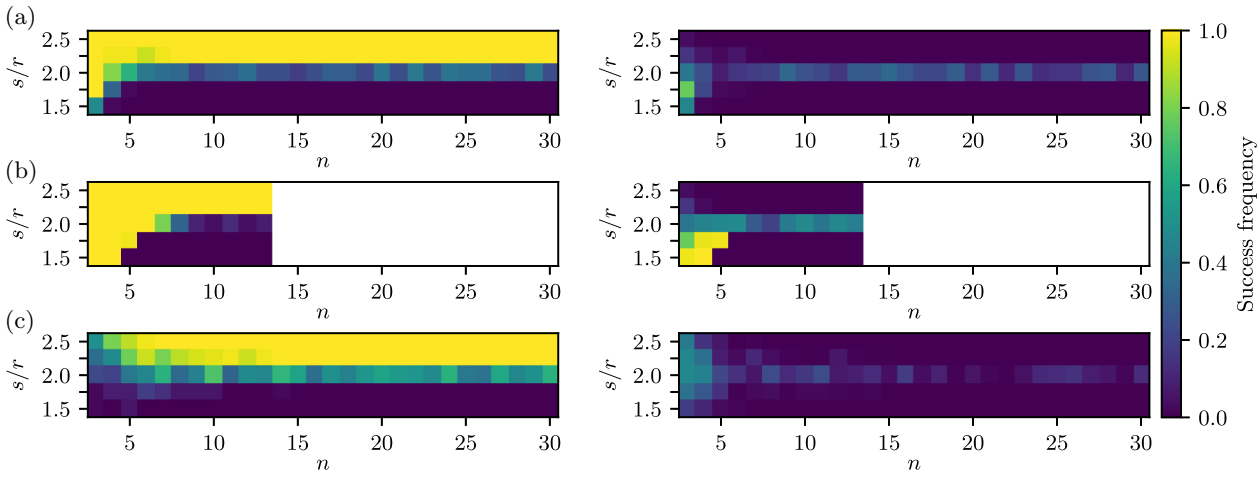


FIG. 6. We select r rows of the Walsh-Hadamard matrix as determined by the considered interactions in the system Hamiltonian, to obtain the partial Walsh-Hadamard matrix $W^{(r\times 4^n)}$. Then, we implement the subsampling of the possible pulse layers by randomly sampling s columns to obtain $W^{(r\times s)}$. The plots in the left column show the success frequency for "$W^{(r\times s)}$ is feasible" (yellow) over 50 samples for $n = 3, \ldots, 30$ and $s/r = 1.5, \ldots, 2.5$. There is a sharp transition at $s/r = 2$ in all cases. The plots in the right column show the difference of our numerical observation compared to Wendel's formula in Eq. (38). (a),(b) The same (a) 2- and (b) 3-local system Hamiltonians as in Figs. 2(a) and 2(b). In (b) we have stopped the numerical experiments at $n = 13$ due to the large amount of all-to-all interactions with locality $k \leq 3$. (c) The same random system Hamiltonian as in Fig. 2(c).

$W^{(r\times s)}$ is independent of the choice of $\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\mathbf{0}\}$. As $r = \mathsf{poly}(n)$ for local Hamiltonians, the relaxed Eq. (PauliLP) can be solved efficiently, and it provides an exact decomposition of the target Hamiltonian. The evolution time $\mathbf{1}^T\boldsymbol{\lambda}$ may, however, not be minimal anymore. The quality of the relaxation can be improved by increasing s, thus expanding the search space. This provides a flexible trade-off between the run-time of Eq. (PauliLP) and the optimality of $\boldsymbol{\lambda}$ [see Figs. 2(a)–2(c)].

B. Clifford conjugation

In Sec. IV A, we have introduced the Hamiltonian-engineering method based on conjugation with π pulses. In this section, we extend this method to a certain set of single-qubit Clifford gates, consisting of $\pi/2$ pulses. This extension allows us to change the interaction terms in the system Hamiltonian, such that the only restriction in engineering a target Hamiltonian comes from the locality of the interactions in the system Hamiltonian H_S. The same restriction also holds for conjugation with arbitrary single-qubit pulses, making Clifford conjugation a powerful Hamiltonian-engineering method. The simplest set of $\pi/2$ pulses with full expressivity are also called the square-root Pauli gates:

$$\sqrt{X} = \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} & 1-\mathrm{i} \\ 1-\mathrm{i} & 1+\mathrm{i} \end{pmatrix}, \quad \sqrt{Y} = \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} & -1-\mathrm{i} \\ 1+\mathrm{i} & 1+\mathrm{i} \end{pmatrix}, \quad \sqrt{Z} = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix}. \tag{39}$$

As displayed in Table II, the linear transformation $S^\dagger PS$ of a Pauli operator with a $\pi/2$ pulse not only flips the sign but also changes the interaction type (rotation axis). The change of the interaction type increases the set of Hamiltonians reachable by Clifford conjugation compared to the Pauli conjugation. In the following, we consider a gate set that behaves very similarly to the conjugation with π pulses:

$$\mathcal{C}_{XY} := \{Z\} \cup \left\{ QD \,\middle|\, Q, D \in \{\sqrt{X}, \sqrt{Y}, \sqrt{X}^\dagger, \sqrt{Y}^\dagger\} \right\}. \tag{40}$$

Likewise, we can define gate sets $\mathcal{C}_{ZY}$ and $\mathcal{C}_{XZ}$, for which conclusions analogous to the following ones may be drawn. For example, the transformation given by the conjugation of interaction terms with the Clifford gates $\sqrt{X}\sqrt{Y}$ (two $\pi/2$ pulses) or $\sqrt{Y}^\dagger\sqrt{X}^\dagger$ changes the interaction type and leaves the sign of the Pauli coefficient unchanged. The signs of the conjugated interaction terms

TABLE II. The conjugation of Pauli operators $P \in \mathsf{P}$ with square-root Pauli gates ($\pi/2$ pulses) $S \in \{\sqrt{X}, \sqrt{Y}, \sqrt{Z}, \sqrt{X}^\dagger, \sqrt{Y}^\dagger, \sqrt{Z}^\dagger\}$.

$S^\dagger PS$	$\sqrt{X}$	$\sqrt{Y}$	$\sqrt{Z}$	$\sqrt{X}^\dagger$	$\sqrt{Y}^\dagger$	$\sqrt{Z}^\dagger$
X	X	Z	$-Y$	X	$-Z$	Y
Y	$-Z$	Y	X	Z	Y	$-X$
Z	Y	$-X$	Z	$-Y$	X	Z

TABLE III. The colored cells show the conjugated interaction terms $S_c^\dagger P_{\boldsymbol{a}} S_c$ for an interaction $P_{\boldsymbol{a}} \in \mathsf{P}$ conjugated by the single-qubit Clifford gate $S_c \in \mathcal{C}_{XY}$. We label an element in $\mathcal{C}_{XY}$ by $c := (p, \boldsymbol{b}) \in \mathbb{F}_3 \times \mathbb{F}_2^2$, where $p \in \mathbb{F}_3$ captures the permutation of the conjugated interaction terms, and $\boldsymbol{b}$ captures the sign flips similar to the Pauli conjugation. This can be easily verified using Table II.

p	$\boldsymbol{b}$	$S_{p,\boldsymbol{b}}$	$P_{(0,0)} = I$	$P_{(1,0)} = X$	$P_{(1,1)} = Y$	$P_{(0,1)} = Z$
0	$(0,0)$	I	I	X	Y	Z
	$(1,0)$	X	I	X	$-Y$	$-Z$
	$(1,1)$	Y	I	$-X$	Y	$-Z$
	$(0,1)$	Z	I	$-X$	$-Y$	Z
1	$(0,0)$	$\sqrt{X}\sqrt{Y}$	I	Z	X	Y
	$(1,0)$	$\sqrt{X}^\dagger\sqrt{Y}$	I	Z	$-X$	$-Y$
	$(1,1)$	$\sqrt{X}^\dagger\sqrt{Y}^\dagger$	I	$-Z$	X	$-Y$
	$(0,1)$	$\sqrt{X}\sqrt{Y}^\dagger$	I	$-Z$	$-X$	Y
2	$(0,0)$	$\sqrt{Y}^\dagger\sqrt{X}^\dagger$	I	Y	Z	X
	$(1,0)$	$\sqrt{Y}\sqrt{X}$	I	Y	$-Z$	$-X$
	$(1,1)$	$\sqrt{Y}\sqrt{X}^\dagger$	I	$-Y$	Z	$-X$
	$(0,1)$	$\sqrt{Y}^\dagger\sqrt{X}$	I	$-Y$	$-Z$	X

depend on the rotation direction, i.e., the placement of "†," of the square-root Pauli gates (see Table III for examples). Therefore, we label an element in $\mathcal{C}_{XY}$ by $c := (p, \boldsymbol{b}) \in \mathbb{F}_3 \times \mathbb{F}_2^2$, where $p \in \mathbb{F}_3$ represents the changes in the interaction type and $\boldsymbol{b} \in \mathbb{F}_2^2$ captures the sign flips similar to the Pauli conjugation. Denote by $\mathcal{C}_{XY}^{\otimes n}$ a string of single-qubit gates on n qubits from the gate set $\mathcal{C}_{XY}$. We label each $S_{\boldsymbol{c}} \in \mathcal{C}_{XY}^{\otimes n}$ with $\boldsymbol{c} = (c_1, \ldots, c_n) = (\boldsymbol{p}, \boldsymbol{b}) \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}$, where $c_i \in \mathbb{F}_3 \times \mathbb{F}_2^2$ represents the single-qubit Clifford gate from $\mathcal{C}_{XY}$ on the ith qubit. In Table III, we show the effect of conjugating an interaction term $P_{\boldsymbol{a}} \in \mathsf{P}$ with $S_c \in \mathcal{C}_{XY}$.

The transformation of H_S with respect to $S_{\boldsymbol{c}} \in \mathcal{C}_{XY}^{\otimes n}$ leads to

$$\begin{aligned} S_{\boldsymbol{c}}^\dagger H_S S_{\boldsymbol{c}} &= \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} J_{\boldsymbol{a}} S_{\boldsymbol{c}}^\dagger P_{\boldsymbol{a}} S_{\boldsymbol{c}} \\ &= \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} (-1)^{\langle \pi_{\boldsymbol{p}}(\boldsymbol{a}), \boldsymbol{b} \rangle} J_{\pi_{\boldsymbol{p}}(\boldsymbol{a})} P_{\boldsymbol{a}}, \end{aligned} \tag{41}$$

with $\boldsymbol{c} = (\boldsymbol{p}, \boldsymbol{b}) \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}$ and the permutation $\pi_{\boldsymbol{p}} : \mathbb{F}_2^{2n} \to \mathbb{F}_2^{2n}$ with $\pi_{\boldsymbol{p}}(\boldsymbol{a}) := \big(\pi_{p_1}(a_1), \ldots, \pi_{p_n}(a_n)\big)$ given by the local permutations $\pi_0, \pi_1, \pi_2 : \mathbb{F}_2^2 \to \mathbb{F}_2^2$ in the two-line notation

$$\begin{aligned} \pi_0 &:= \begin{pmatrix} (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \end{pmatrix}, \\ \pi_1 &:= \begin{pmatrix} (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}, \\ \pi_2 &:= \begin{pmatrix} (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (0,1) & (1,0) & (1,1) \end{pmatrix}. \end{aligned} \tag{42}$$

We want to find a decomposition of H_T with $\lambda_{\boldsymbol{c}} \geq 0$ such that

$$H_T = \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} A_{\boldsymbol{a}} P_{\boldsymbol{a}} = \sum_{\boldsymbol{c} \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}} \lambda_{\boldsymbol{c}} S_{\boldsymbol{c}}^\dagger H_S S_{\boldsymbol{c}}. \tag{43}$$

We define the matrix $W(\boldsymbol{J}) \in \mathbb{R}^{(4^n-1)\times 12^n}$ with entries

$$W(\boldsymbol{J})_{\boldsymbol{a}\boldsymbol{c}} := (-1)^{\langle \pi_{\boldsymbol{p}}(\boldsymbol{a}), \boldsymbol{b} \rangle} J_{\pi_{\boldsymbol{p}}(\boldsymbol{a})}, \tag{44}$$

for $\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}$, excluding the identity term $P_{\boldsymbol{a}} = I^{\otimes n}$ with $\boldsymbol{a} = \boldsymbol{0}$, and $\boldsymbol{c} \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}$. Similarly to the Pauli conjugation we define the submatrix $W(\boldsymbol{J})^{(r\times s)} \in \mathbb{R}^{r\times s}$ with entries given in Eq. (44) for r row indices $\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}$ and s column indices $\boldsymbol{c} \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}$. Up to the permutation of $\boldsymbol{J}$, the same Walsh-Hadamard matrix structure as in $W^{(r\times s)}$ from Sec. IV A is present in $W(\boldsymbol{J})^{(r\times s)}$. In terms of the matrix $W(\boldsymbol{J})$ it follows that

$$H_T = \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} \sum_{\boldsymbol{c} \in \mathbb{F}_3^n \times \mathbb{F}_2^{2n}} \lambda_{\boldsymbol{c}} W(\boldsymbol{J})_{\boldsymbol{a}\boldsymbol{c}} P_{\boldsymbol{a}}. \tag{45}$$

Due to the permutation in Eq. (44) we are no longer restricted by the nonzero coefficients as in the Pauli conjugation. From Eq. (45) it follows that for each $A_{\boldsymbol{a}} \neq 0$ there has to be at least one $J_{\hat{\boldsymbol{a}}} \neq 0$ such that $\mathrm{supp}(\hat{\boldsymbol{a}}) = \mathrm{supp}(\boldsymbol{a})$. Therefore, we define

$$\begin{aligned} \mathrm{suppnz}(\boldsymbol{J}) := \big\{ \boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\} \,\big|\, \exists \hat{\boldsymbol{a}} \in \mathrm{nz}(\boldsymbol{J}), \\ \mathrm{supp}(\hat{\boldsymbol{a}}) = \mathrm{supp}(\boldsymbol{a}) \big\} \end{aligned} \tag{46}$$

and require that the Pauli coefficients satisfy

$$\mathrm{nz}(\boldsymbol{A}) \subseteq \mathrm{suppnz}(\boldsymbol{J}). \tag{47}$$

In physical terms, this means that we are only restricted by the locality of the interactions in the system Hamiltonian H_S and have full flexibility in the kind of interactions $P_{\boldsymbol{a}}$ and the interaction strength $A_{\boldsymbol{a}}$. In addition to the restriction $\boldsymbol{a} \in \mathrm{suppnz}(\boldsymbol{J})$, given by the system Hamiltonian, we can also consider a restricted set of conjugating Clifford strings

$\boldsymbol{c} \in \mathcal{F} \subseteq \mathbb{F}_3^n \times \mathbb{F}_2^{2n}$, as long as there still exists a solution. Then, the restricted Eq. (45) can be reformulated as

$$A_{\boldsymbol{a}} = \sum_{\boldsymbol{c}\in\mathcal{F}} W(\boldsymbol{J})^{(r\times s)}_{\boldsymbol{ac}} \lambda_{\boldsymbol{c}} \quad \forall \boldsymbol{a} \in \mathrm{suppnz}(\boldsymbol{J}), \tag{48}$$

where $r := |\mathrm{suppnz}(\boldsymbol{J})|$ denotes the number interactions that can be generated and $s = |\mathcal{F}|$ denotes the number of Clifford strings considered. The constraint in Eq. (LP) then reads $\boldsymbol{A} = W(\boldsymbol{J})^{(r\times s)}\boldsymbol{\lambda}$, with $\boldsymbol{A}, \boldsymbol{J} \in \mathbb{R}^r$ restricted to $\boldsymbol{a} \in \mathrm{suppnz}(\boldsymbol{J})$. With that, we define the following LP:

$$\begin{aligned} &\text{minimize} && \mathbf{1}^T\boldsymbol{\lambda} \\ &\text{subject to} && W(\boldsymbol{J})^{(r\times s)}\boldsymbol{\lambda} = \boldsymbol{A},\ \boldsymbol{\lambda} \in \mathbb{R}^s_{\geq 0}. \end{aligned} \tag{CliffLP}$$

There always exists a feasible solution of Eq. (CliffLP) with $W(\boldsymbol{J})^{(r\times 12^n)}$, which follows similarly to Corollary B1 from the Walsh-Hadamard structure in $W(\boldsymbol{J})^{(r\times 12^n)}$. In general, we have $s \leq 12^n$ and $r \leq 4^n - 1$. In the next section, we propose a simple and efficient relaxation of Eq. (CliffLP), where $s \geq 2r$ can be chosen for arbitrary $r = |\mathrm{suppnz}(\boldsymbol{J})|$.

1. Efficient relaxation

The size of Eq. (CliffLP) scales as $\mathrm{O}(12^n)$ if all possible Clifford layers are considered. Therefore, a relaxation is necessary to solve this LP in practice. Analogous to the Pauli-conjugation method (Sec. IV A 1), sampling $s \geq 2r$ columns at random from $W(\boldsymbol{J})^{(r\times 12^n)}$ results in a feasible matrix $W(\boldsymbol{J})^{(r\times s)}$ with high probability. As in Sec. IV A 1, we test this statement numerically by checking the feasibility conditions of Proposition 1 for random submatrices.

Numerical observation 2 (Fig. 7). Let $W(\boldsymbol{J})^{(r\times 12^n)}$ be a matrix with $r = |\mathrm{suppnz}(\boldsymbol{J})|$ and entries as in Eq. (44), and let $\boldsymbol{w}$ be a r-dimensional random vector drawn uniformly from $\mathrm{col}(W(\boldsymbol{J})^{(r\times 12^n)})$. Then, we numerically observe that Wendel's statement in Eq. (38) approximately holds, i.e., the random submatrix $W(\boldsymbol{J})^{(r\times s)}$ is feasible with high probability provided that we take $s \geq 2r$ samples of $\boldsymbol{w}$.

From the definition of feasibility, we know that Eq. (CliffLP) always has a solution for arbitrary $\boldsymbol{A}$, i.e., arbitrary $\boldsymbol{J}, \boldsymbol{A}$ such that $\mathrm{nz}(\boldsymbol{A}) \subseteq \mathrm{suppnz}(\boldsymbol{J})$. Furthermore, the quality of the relaxation can be improved by increasing s, which leads to an expanded search space [see Figs. 2(d)–2(f)]. This again provides a flexible trade-off between the run-time of Eq. (CliffLP) and the optimality of $\boldsymbol{\lambda}$.

C. Hamiltonian engineering with unknown Hamiltonians

In practice, sometimes not all coupling coefficients $J_{\boldsymbol{a}}$ in the system Hamiltonian might be known; e.g., when an experiment aims to realize two-body interactions but also acquires unwanted three-body terms of unknown strength. Such system Hamiltonians with unknown or only partially known coupling strengths can still be used for engineering based on the Pauli-conjugation method from Sec. IV A. Solving Eq. (PauliLP) for an $\boldsymbol{M} \in \mathbb{R}^r$ yields the target

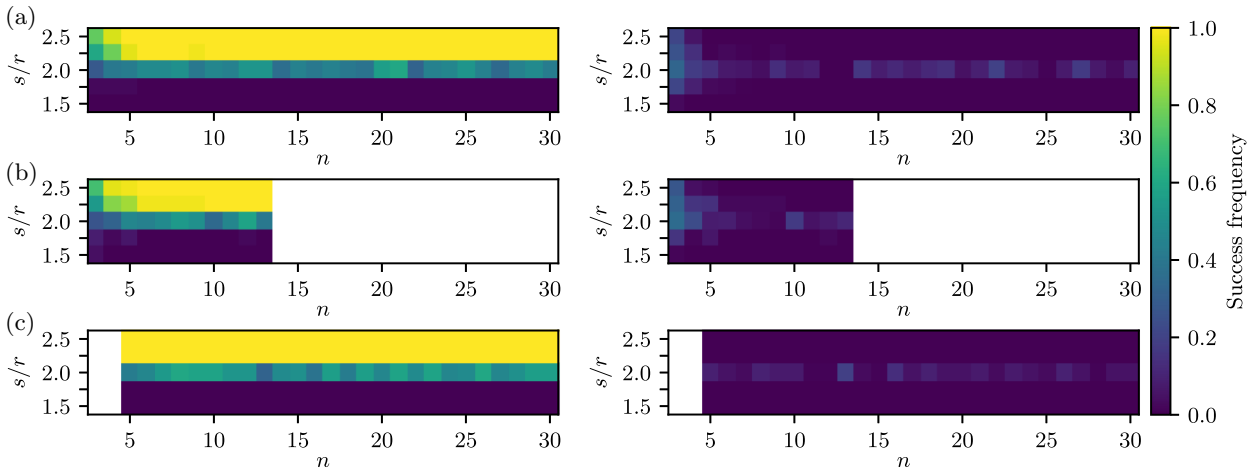


FIG. 7. We select r rows of $W(\boldsymbol{J})$, determined by the considered interactions in the system Hamiltonian, to obtain the partial matrix $W(\boldsymbol{J})^{(r\times 12^n)}$. Then, we implement the subsampling of the possible pulse layers by randomly sampling s columns to obtain $W(\boldsymbol{J})^{(r\times s)}$. The plots in the left column show the success frequency over 50 samples for "$W(\boldsymbol{J})^{(r\times s)}$ is feasible" (yellow) for each $n = 3, \ldots, 30$ and $s/r = 1.5, \ldots, 2.5$. There is a sharp transition at $s/r = 2$ in all cases. The plots in the right column show the difference of our numerical observation compared to Wendel's statement in Eq. (38). (a),(b) The same (a) 2- and (b) 3-local system Hamiltonians as in Figs. 2(d) and 2(e), respectively. In (b) we have stopped the numerical experiments at $n = 13$ due to the large amount of all-to-all interactions with locality $k \leq 3$. (c) The same random 5-local system Hamiltonian as in Fig. 2 (f).

Hamiltonian

$$\sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} M_{\boldsymbol{a}} J_{\boldsymbol{a}} P_{\boldsymbol{a}} = H_T. \tag{49}$$

The potentially unknown coefficients $J_{\boldsymbol{a}}$ are modified by an element-wise multiplication $\boldsymbol{M} \odot \boldsymbol{J}$. Interesting choices for $M_{\boldsymbol{a}}$ are -1 and 0, inverting the sign of or canceling the interaction term $P_{\boldsymbol{a}}$, respectively, without the knowledge of $J_{\boldsymbol{a}}$. For each term in the system Hamiltonian, a different $M_{\boldsymbol{a}}$ can be chosen. Therefore, engineering known terms $M_{\boldsymbol{a}} = A_{\boldsymbol{a}}/J_{\boldsymbol{a}}$ in the Hamiltonian while canceling or inverting other (potentially unknown) terms $M_{\boldsymbol{a}} = 0$ or $M_{\boldsymbol{a}} = -1$ is possible. Using such a $\boldsymbol{M}$ in Eq. (PauliLP) with dummy values for the unknown $J_{\boldsymbol{a}}$ inverts the signs or cancels the interactions.

In Sec. III B 1 we have applied the approach to cancel unknown three-body terms in a 2D lattice Hamiltonian.

V. ERROR ROBUSTNESS AND MITIGATION TECHNIQUES

To successfully apply the Pauli or Clifford conjugation in practice it is necessary to make the resulting pulse sequences robust against dominant errors. In this section, we provide mitigation techniques for our efficient conjugation methods, which come with little overhead. The simultaneous action of the single-qubit pulse and system Hamiltonian is called the finite-pulse-time error. It has been shown that this error is detrimental to approaches similar to ours [65]. Therefore, our focus lies on the error due to a finite single-qubit pulse duration. Furthermore, we provide a modification to combine our Clifford method with robust composite pulses, making it robust against many different errors occurring in experiments. We achieve robustness against finite-pulse-time effects and rotation-angle errors using a similar approach as in the work of Votto *et al.* [18]. Their study focuses on specifically designed π-pulse sequences (i.e., Pauli gates), so-called Walsh sequences, for engineering XY Hamiltonians. In our work, we generalize their approach to general local system Hamiltonians and arbitrary single-qubit pulses.

AHT is a well-known approach in NMR that allows us to investigate the dynamics under a time-dependent Hamiltonian by approximating it with that of a time-dependent Hamiltonian [66,67]. This approximation is done by a low-order Magnus expansion [68]. We utilize AHT to investigate the error due to a finite single-qubit pulse time. We will heavily use the fact that the error term in the average Hamiltonian has the same locality as the system Hamiltonian.

In the following, we give an explicit form of the finite-pulse-time error in the average Hamiltonian when interleaving a system Hamiltonian with arbitrarily many layers of single-qubit pulses. For the Clifford conjugation method, we find that the finite-pulse-time error can be exactly cancelled by a slight modification of Eq. (CliffLP). This general investigation enables us to also mitigate the finite-pulse-time error in combination with other pulse errors by replacing the single-qubit Clifford pulses in our Clifford conjugation method with robust composite pulses. The latter are designed to compensate for experimental errors by implementing a gate with a specific pulse sequence. In this way, dominant error sources can be cancelled or suppressed, which includes rotation-angle errors (Rabi-frequency errors), off-resonance errors [49, 50], phase errors [69], pulse-shape errors [70–72], or non-stationary, non-Markovian noise [73] or crosstalk [74]. If needed, remaining errors might then be mitigated using software-based methods [75].

For the Pauli-conjugation method, the finite-pulse-time error can only be partially mitigated by modifying Eq. (PauliLP). However, the freedom in choosing the rotation direction in the π pulse can be leveraged in addition to the modified Eq. (PauliLP) to completely cancel the finite pulse time error term in the average Hamiltonian. Moreover, this simultaneously cancels first-order effects of rotation-angle errors in the single-qubit pulses.

We introduce basic concepts required for the following sections in Sec. V A. Then, in Sec. V B, we start with the investigation of the finite-pulse-time error when conjugating a system Hamiltonian with arbitrary many general single-qubit pulses. In Sec. V C, we apply the general results to present our robust Clifford-conjugation methods. Finally, in Sec. V D, we present our robust Pauli-conjugation method.

A. Preliminaries

For our robust methods, we require two well-known approaches to approximating the time evolution under non-commuting Hamiltonians. First, we present general Trotter product formulas to approximate the time evolution under a linear combination of time-independent Hamiltonians. Recently, the performance of the Suzuki-Trotter approximation has been greatly improved [76]. Second, we present the Magnus expansion to approximate the time evolution under a time-dependent Hamiltonian by a time-independent Hamiltonian.

1. Product formula

The time evolution under a Hamiltonian $H = \sum_{i=1}^{L} H_i$ can generally be approximated by a *product formula*,

$$\mathrm{e}^{-\mathrm{i}tH} \approx \mathrm{e}^{-\mathrm{i}\alpha_q H_{i_q}} \dots \mathrm{e}^{-\mathrm{i}\alpha_1 H_{i_1}}, \tag{50}$$

with the number of evolution steps q and $i_1, \dots, i_q \in [L]$. We call a product formula *deterministic* if $i_1, \dots, i_q$ can be found by a deterministic algorithm. For our methods we do not require that $\alpha_1, \dots, \alpha_q \in \mathbb{R}$ are chosen deterministically. The best known deterministic product formulas

are the Suzuki-Trotter formulas [77,78]. The first- and second-order approximations are given by

$$\mathrm{e}^{-\mathrm{i}tH} \approx \left(\prod_{i=1}^{L\rightarrow} \mathrm{e}^{-\mathrm{i}\frac{t}{n_{\mathrm{Tro}}}H_i}\right)^{n_{\mathrm{Tro}}} =: S_1(t/n_{\mathrm{Tro}})^{n_{\mathrm{Tro}}} \tag{51}$$

and

$$\mathrm{e}^{-\mathrm{i}tH} \approx \left(\prod_{i=1}^{L\rightarrow} \mathrm{e}^{-\mathrm{i}\frac{t}{2n_{\mathrm{Tro}}}H_i} \prod_{i=1}^{L\rightarrow} \mathrm{e}^{-\mathrm{i}\frac{t}{2n_{\mathrm{Tro}}}H_i}\right)^{n_{\mathrm{Tro}}} =: S_2(t/n_{\mathrm{Tro}})^{n_{\mathrm{Tro}}}, \tag{52}$$

respectively, with the *number of Trotter cycles* n_{Tro}. The number of Trotter cycles can be increased to improve the accuracy. The $2k$th-order Suzuki-Trotter formula for $k > 1$ is defined recursively by

$$S_{2k}(t) := S_{2k-2}(u_k t)^2 S_{2k-2}((1-4u_k)t) S_{2k-2}(u_k t)^2, \tag{53}$$

with $u_k := (4 - 4^{(2k-1)^{-1}})^{-1}$. For a $2k$th-order Suzuki-Trotter formula the approximation error in the spectral norm is bounded by

$$\|S_{2k}(t/n_{\mathrm{Tro}})^{n_{\mathrm{Tro}}} - \mathrm{e}^{-\mathrm{i}tH}\| \leq \mathrm{O}\left(\left(t\sum_{i=1}^{L}\|H_i\|\right)^{2k+1} {n_{\mathrm{Tro}}}^{-2k}\right), \tag{54}$$

with the spectral norm $\|H_i\|$ [79].

2. Average Hamiltonian theory and the Magnus expansion

Given a time-dependent Hamiltonian $H(t)$, it is sometimes useful to consider a time-independent effective or *average Hamiltonian* H_{av} satisfying

$$\mathcal{U}(T) \approx \mathrm{e}^{-\mathrm{i}TH_{\mathrm{av}}}, \tag{55}$$

where $\mathcal{U}(T)$ is the time-evolution operator defined by the differential equation

$$\frac{d\mathcal{U}(t)}{dt} = -\mathrm{i}H(t)\mathcal{U}(t), \quad \text{and} \quad \mathcal{U}(0) = \mathbb{1}. \tag{56}$$

This average Hamiltonian can be expressed by the Magnus expansion as follows:

$$H_{\mathrm{AV}} = H_{\mathrm{av}}^{(1)} + H_{\mathrm{av}}^{(2)} + \ldots, \tag{57}$$

where the first- and second-order terms are explicitly given by

$$H_{\mathrm{av}}^{(1)} := \frac{1}{T}\int_0^T H(\tau)\mathrm{d}\tau \tag{58}$$

and

$$H_{\mathrm{av}}^{(2)} := \frac{1}{2\mathrm{i}T}\int_0^T\int_0^\tau [H(\tau), H(\tau')]\mathrm{d}\tau'\mathrm{d}\tau. \tag{59}$$

The Magnus expansion converges if $\int_0^T ||H(\tau)||\mathrm{d}\tau < \pi$ [80]. As a rule of thumb, the Magnus expansion converges rapidly if

$$\max_{\tau\in[0,T]} ||H(\tau)||T \ll 1, \tag{60}$$

where the spectral norm is used [81]. Throughout this work we only consider the first-order approximation and write $H_{\mathrm{av}} = H_{\mathrm{av}}^{(1)}$.

B. General robust conjugation method

We start with a general discussion of the finite-pulse-time error when conjugating a system Hamiltonian H_S with multiple layers of general single-qubit pulses. To this end, we first define the Hamiltonian of a general single-qubit pulse layer and the operators relevant for the following discussion. In the following we define different compositions of single-qubit pulse layers, which allow a clear and compact presentation of the main results. In particular, we define a single pulse layer, multiple consecutive pulse layers, which we call a pulse block, and a partial pulse block.

Definition 3. Let $\mathcal{G} \subset \mathrm{Herm}(\mathbb{C}^2)$ be a set of *single-qubit pulse generators*. A *single-qubit pulse layer* is given by a *local generator* h_i chosen from $\mathcal{G}$ for each qubit $i \in [n]$, the *(single-qubit) rotation directions* $\boldsymbol{s} \in \mathbb{F}_2^n$ and the *finite pulse time* $t_p > 0$. The Hamiltonian generating the single-qubit pulse layer on n qubits is given by

$$H(t_p, \boldsymbol{s}, \boldsymbol{h}) := \frac{1}{t_p}\sum_{i=1}^{n}(-1)^{s_i}h_i, \tag{61}$$

with $\boldsymbol{h} := (h_1, \ldots, h_n)$. The single-qubit pulse layer Hamiltonian is completely specified by the tuple $c := (t_p, \boldsymbol{s}, \boldsymbol{h})$, and we write $H_c := H(t_p, \boldsymbol{s}, \boldsymbol{h})$. This generates the evolution operator for one single-qubit pulse layer

$$\boldsymbol{S}_c(t) := \mathrm{e}^{-\mathrm{i}tH_c}, \quad \text{and} \quad \boldsymbol{S}_c := \boldsymbol{S}_c(t_p). \tag{62}$$

More generally, we consider a sequence of n_L single-qubit pulse layers and introduce the layer index $\ell \in [n_L]$. The evolution for the ℓth single-qubit pulse layer is specified by the tuple $c^{(\ell)} := (t_p^{(\ell)}, \boldsymbol{s}^{(\ell)}, \boldsymbol{h}^{(\ell)})$, and we define the tuple $\boldsymbol{c} := (c^{(1)}, \ldots, c^{(n_L)})$. Moreover, we define the *single-qubit*

pulse block and the *partial single-qubit pulse block* as

$$\boldsymbol{S}_{\boldsymbol{c}} := \overrightarrow{\prod_{\ell=1}^{n_L}} \boldsymbol{S}_{c^{(\ell)}} \tag{63}$$

and

$$\boldsymbol{S}_{\boldsymbol{c}^{\geq D}} := \begin{cases} \overrightarrow{\prod_{\ell=D}^{n_L}} \boldsymbol{S}_{c^{(\ell)}}, & 1 \leq D \leq n_L, \\ \mathbb{1}, & \text{otherwise}, \end{cases} \tag{64}$$

respectively. Similarly, we define the pulse time for the single-qubit pulse block $T_p := \sum_{\ell=1}^{n_L} t_p^{(\ell)}$ and for the partial single-qubit pulse block $T_p^{\leq D} := \sum_{\ell=1}^{D} t_p^{(\ell)}$.

To illustrate Definition 3 we provide the Hamiltonian for the single-qubit Pauli pulses from Sec. IV A. The Hamiltonian for a single-qubit Pauli pulse layer is given by the π-pulse time t_p, an arbitrary rotation direction $\boldsymbol{s} \in \mathbb{F}_2^n$, and the generators $h_i = (\pi/2)P_i$, with the Pauli operator P_i given by $P_{\boldsymbol{b}_i}$ from Eq. (32).

Recall from Sec. IV that ideally, we would like to implement the conjugation $\boldsymbol{S}_{\boldsymbol{c}}^\dagger \mathrm{e}^{-\mathrm{i}t\lambda_{\boldsymbol{c}} H_S} \boldsymbol{S}_{\boldsymbol{c}} = \mathrm{e}^{-\mathrm{i}t\lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}}}$, with the system Hamiltonian H_S and the *relative evolution time* $\lambda_{\boldsymbol{c}}$ associated with a single-qubit pulse block $\boldsymbol{S}_{\boldsymbol{c}}$. However, due to the finite duration of the single-qubit pulse and the always-on system Hamiltonian, we get the time evolution of the single-qubit pulse blocks,

$$\boldsymbol{S}_{\mathrm{err},\boldsymbol{c}} = \overrightarrow{\prod_{\ell=1}^{n_L}} \mathrm{e}^{-\mathrm{i}t_p^{(\ell)}(H_S - H_{c^{(\ell)}})} \tag{65}$$

and

$$\boldsymbol{S}'_{\mathrm{err},\boldsymbol{c}} = \overrightarrow{\prod_{\ell=1}^{n_L}} \mathrm{e}^{-\mathrm{i}t_p^{(\ell)}(H_S + H_{c^{(\ell)}})}, \tag{66}$$

respectively, with the finite duration of the ℓth single-qubit pulse layer $t_p^{(\ell)} > 0$. In the absence of H_S, these two operators are exact inverses of each other. Then, the resulting evolution block with the finite-pulse-time error has the form

$$U(t\lambda_{\boldsymbol{c}}) := \boldsymbol{S}'_{\mathrm{err},\boldsymbol{c}} \mathrm{e}^{-\mathrm{i}t\lambda_{\boldsymbol{c}} H_S} \boldsymbol{S}_{\mathrm{err},\boldsymbol{c}}, \tag{67}$$

similar to as in Eq. (17), and is depicted in Fig. 8.

First, we provide the average Hamiltonian for one conjugation with a single-qubit pulse block to investigate the effect of the finite pulse duration.

Lemma 4. The approximation of $U(t\lambda_{\boldsymbol{c}})$ in first-order Magnus expansion is given by $\mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)}$ with

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}, \tag{68}$$

where

$$H_{\mathrm{err},\boldsymbol{c}} := 2 \sum_{\ell=1}^{n_L} \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^\dagger \int_0^{t_p^{(\ell)}} \boldsymbol{S}_{c^{(\ell)}}^\dagger(t) H_S \boldsymbol{S}_{c^{(\ell)}}(t) \mathrm{d}t \, \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}. \tag{69}$$

The average Hamiltonian $H_{\mathrm{av},\boldsymbol{c}}(t)$ has the same locality as H_S. Moreover, the error for truncating the Magnus expansion after the first order can be bounded in the spectral norm by

$$\| U(t\lambda_{\boldsymbol{c}}) - \mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)} \| \leq \mathrm{O}((2T_p + t\lambda_{\boldsymbol{c}})^2 \|H_S\|^2). \tag{70}$$

The proof of Lemma 4 can be found in Appendix D.

Our goal is to formulate an LP that cancels the finite-pulse-time error in the average Hamiltonian (in first order). To this end, we define the matrices capturing the effect of the ideal dynamics $\boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}}$ and the finite-pulse-time effect $H_{\mathrm{err},\boldsymbol{c}}$. The ideal dynamics are captured by the matrix $\mathcal{W}(\boldsymbol{J})^{(r\times s)} \in \mathbb{R}^{r\times s}$ with the same elements as in Eq. (28).

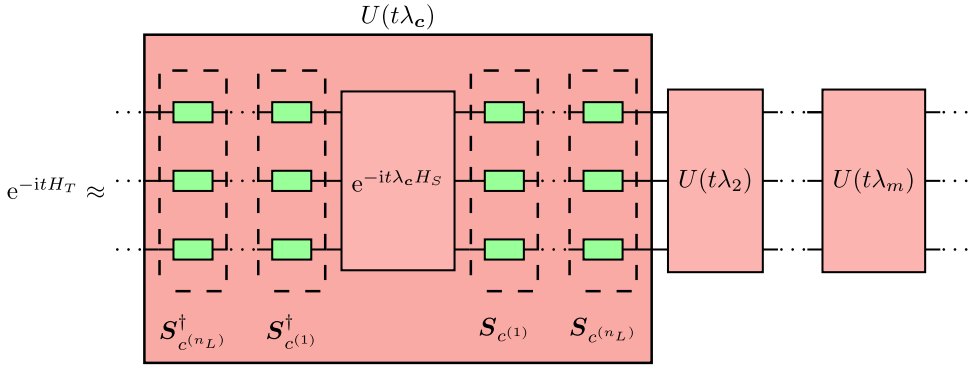


FIG. 8. An exemplary quantum circuit for approximating the target evolution. We only implement simple single-qubit pulses in the presence of an always-on system Hamiltonian H_S.

The effect of the finite-pulse-time error is captured by the matrix $E(\boldsymbol{J})^{(r\times s)} \in \mathbb{R}^{r\times s}$ with the elements

$$E(\boldsymbol{J})^{(r\times s)}_{\boldsymbol{ac}} := \frac{1}{2^n}\,\mathrm{Tr}\left(P_{\boldsymbol{a}} H_{\mathrm{err},\boldsymbol{c}}\right) \tag{71}$$

and can be efficiently calculated for any local system Hamiltonian H_S (see Lemma 4 in Appendix D). Let $\mathcal{W}(\boldsymbol{J})^{(r\times s)}$ be feasible, i.e., let $\sum_{\boldsymbol{c}} \lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}}$ be able to modify all interaction terms with the same locality as the interactions in H_S. Then, the LP

$$\begin{aligned} &\text{minimize} && \mathbf{1}^T\boldsymbol{\lambda} \\ &\text{subject to} && \mathcal{W}(\boldsymbol{J})^{(r\times s)}\boldsymbol{\lambda} + E(\boldsymbol{J})^{(r\times s)}\mathbf{1} = \boldsymbol{A}, \quad \text{(robustLP)} \\ &&& \boldsymbol{\lambda} \in \mathbb{R}^s_{\geq 0} \end{aligned}$$

always has a solution for any target interaction strength $\boldsymbol{A} \in \mathbb{R}^r$. This follows directly from Definition 1. Now, we are ready to state the main result of this section.

Theorem 5. The target time evolution e^{-itH_T} can be efficiently approximated by a deterministic product formula implementing a product of $U(t\lambda_{\boldsymbol{c}})$, with $\lambda_{\boldsymbol{c}}$ and the corresponding pulse block $\boldsymbol{S}_{\boldsymbol{c}}$ from Eq. (robustLP). Moreover, this approximation is robust against the finite-pulse-time effect. The only approximation errors are given by the ones from the Magnus approximation (Lemma 4) and the approximation error from the deterministic product formula.

The proof of Theorem 5 can be found in Appendix D.

Remark 1. The Trotter approximation error can be bounded as in Eq. (54) and depends on the evolution time $t\sum_{\boldsymbol{c}} \lambda_{\boldsymbol{c}}$. The approximation error for truncating the Magnus expansion increases with the finite pulse duration [see Eq. (70)]. Note that the truncation error bound in Eq. (70) is not tight and might be improved [82].

Remark 2. Equation (robustLP) is efficiently solvable with the relaxation from Sec. IV B 1 if $s \geq 2r$, where r is the number of interaction terms that can be modified. To approximate the evolution under the target Hamiltonian H_T with n_{Tro} Trotter cycles we have to implement at most $n_{\mathrm{Tro}}s$ evolution blocks $U(t\lambda_{\boldsymbol{c}})$, even if $\lambda_{\boldsymbol{c}} = 0$, since the finite-pulse-time errors for all single-qubit pulse layers are taken into account in Eq. (robustLP). Each evolution block contains $2n_L$ single-qubit pulse layers (possibly being part of a robust composite-pulse sequence). Then, the total number of single-qubit pulse layers is at most $2n_L n_{\mathrm{Tro}} s$. The number of single-qubit pulse layers can be reduced from $2n_L n_{\mathrm{Tro}} s$ to $\approx 2n_L n_{\mathrm{Tro}} r$ by formulating a MILP, which we explain in Sec. V E.

To summarize, our efficient relaxation plays a central role in solving the LP formulation Hamiltonian engineering problems.

C. Robust Clifford-conjugation method

In this section, we leverage the general robust conjugation method to robustify the Clifford conjugation method. We have two single-qubit pulse layers $n_L = 2$ of $\pi/2$ pulses, splitting each Pauli π pulse in the gate set $\mathcal{C}_{XY}$ into two $\pi/2$ pulses. Therefore, the ideal evolution in Eq. (robustLP) is captured by the matrix $\mathcal{W}(\boldsymbol{J})^{(r\times s)} = W(\boldsymbol{J})^{(r\times s)}$.

Definition 3 (Single-qubit Clifford pulse block). The *single-qubit Clifford pulse block* $\boldsymbol{S}_{\boldsymbol{c}}$ is specified by the tuple $\boldsymbol{c} = ((t_p^{(1)}, \boldsymbol{s}^{(1)}, \boldsymbol{h}^{(1)}), (t_p^{(2)}, \boldsymbol{s}^{(2)}, \boldsymbol{h}^{(2)}))$, with the finite pulse times of one $\pi/2$ pulse, and we have $t_p^{(1)} = t_p^{(2)}$. The rotation direction $\boldsymbol{s}^{(\ell)}$ is fixed, and we set $s_i^{(\ell)} = 1$ if on the ith qubit and the ℓ layer we have $\sqrt{(\cdot)}^\dagger$ and $s_i^{(\ell)} = 0$ if we have $\sqrt{(\cdot)}$. The generators are $h_i = (\pi/4)P_i^{(\ell)}$, with the Pauli operators from the square-root Pauli gates $P_i^{(\ell)}$.

The Clifford method can be made robust against the finite-pulse-time error by direct application of Theorem 5. Our results for the efficient relaxation of the Clifford conjugation method in Sec. IV B 1 ensure that Eq. (robustLP) is feasible and that all interaction terms with the same locality as the interactions in H_S can be modified.

Corollary 1 (Robustness against finite-pulse-time errors). Let there be two layers of $\pi/2$ pulses, representing the single-qubit pulses for the Clifford conjugation as in Definition 4. Then, the finite-pulse-time effect can be suppressed using Theorem 5.

The Clifford-conjugation method can also be made robust by taking advantage of the rich field of robust composite pulses. The $\pi/2$ pulses in each of the two layers in the Clifford-conjugation method can be made robust by replacing each pulse with robust composite pulses of length $n_L/2$. Then, the finite-pulse-time error is different but can still be corrected using Theorem 5.

Corollary 2 (Robustness against pulse errors). Let there be a robust composite-pulse sequence for $\pi/2$ pulses of length $n_L/2$. Replacing each $\pi/2$ pulse in the Clifford-conjugation method by such robust composite pulses yields n_L single-qubit pulse layers. Then, the finite-pulse-time effect is different from Corollary 1 but can still be suppressed by Theorem 5.

The ability to modify any interaction with the same locality as H_S and the combination with robust composite pulses makes our Clifford conjugation robust against a wide range of different errors in experiments. Note that composite pulses with a short overall duration are more beneficial due to the faster convergence of the Magnus expansion in Eq. (60). In Sec. III B 2 we have combined the

SCROFULOUS pulses [58] and the SCORBUTUS pulses [50] with the robust Clifford conjugation to implement arbitrary Heisenberg Hamiltonians.

D. Robust Pauli-conjugation method

The Pauli conjugation can only change the nonzero interaction strengths in the system Hamiltonian but not modify the type of interactions. Consequently, the general robust conjugation method is not directly applicable. Therefore, we generalize the robustness conditions of Votto *et al.* [18] to arbitrary local Hamiltonians.

Definition 4 (Single-qubit Pauli pulse layer). The *single-qubit Pauli pulse layer* $\boldsymbol{S_c}$ is specified by the tuple $\boldsymbol{c} = (t_p, \boldsymbol{s}, \boldsymbol{h})$, with the finite pulse time t_p of one π pulse. The rotation direction is $\boldsymbol{s} \in \mathbb{F}_2^n$ and can be chosen freely for π pulses. The generators are $h_i = (\pi/2)P_i$, with the Pauli operator P_i given by $P_{\boldsymbol{b}_i}$ from Eq. (33).

For the Pauli conjugation, $U(t\lambda_{\boldsymbol{c}})$ simplifies to Eq. (17). Next, we provide the average Hamiltonian for the conjugation of the system Hamiltonian with a single-qubit Pauli pulse layer to investigate the effect of the finite pulse time.

Lemma 5. Consider all labels $\boldsymbol{a} \in \mathbb{F}_2^{2n}$ for the nonzero interactions $J_{\boldsymbol{a}} \neq 0$ in the system Hamiltonian. Then, the approximation of $U(t\lambda_{\boldsymbol{c}})$ for the Pauli conjugation in the first-order Magnus expansion is given by $\mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)}$ with

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^{\dagger} H_S \boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}, \tag{72}$$

where

$$H_{\mathrm{err},\boldsymbol{c}} = \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} \left(J_{\boldsymbol{a}} E_{\boldsymbol{a},\boldsymbol{c}}^{(r\times s)} P_{\boldsymbol{a}} + R_{\boldsymbol{a},\boldsymbol{c}} \right). \tag{73}$$

We call the first term in Eq. (73) the *interaction term*, with

$$E_{\boldsymbol{a},\boldsymbol{c}}^{(r\times s)} := \frac{4t_p}{\pi} \int_0^{\frac{\pi}{2}} \left(\prod_{i \in \mathrm{supp}(\boldsymbol{a})} (\cos^2(\theta) + (-1)^{\langle \boldsymbol{a}_i, \boldsymbol{b}_i \rangle} \sin^2(\theta)) \right) \mathrm{d}\theta, \tag{74}$$

which we collect as entries of the matrix $E^{(r\times s)} \in \mathbb{R}^{r\times s}$. We call the second term $R_{\boldsymbol{a},\boldsymbol{c}}$ the *rest term*, and it is proportional to $(-1)^{\boldsymbol{e_a}\cdot\boldsymbol{s}}$, with $\boldsymbol{s} \in \mathbb{F}_2^n$ representing the chosen rotation direction of the π pulses and $\boldsymbol{e_a} \in \mathbb{F}_2^n$ encodes the sign flips due to the finite-pulse-time error such that $e_{\boldsymbol{a},i} = 0$ for all $i \notin \mathrm{supp}(\boldsymbol{a})$.

The proof of Lemma 6 can be found in Appendix E. We define the LP similarly to Eq. (robustLP):

$$\begin{aligned} \text{minimize} \quad & \mathbf{1}^T \boldsymbol{\lambda} \\ \text{subject to} \quad & W^{(r\times s)} \boldsymbol{\lambda} + E^{(r\times s)} \mathbf{1} = \boldsymbol{M}, \\ & \boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^s, \end{aligned} \tag{robustPauliLP}$$

with $W^{(r\times s)} \in \mathbb{R}^{r\times s}$ and $\boldsymbol{M} \in \mathbb{R}^r$ from Sec. IV A and $E^{(r\times s)} \in \mathbb{R}^{r\times s}$ from Lemma 6. Similar to Theorem 5 we can cancel the effect of the interaction terms in Eq. (73) by implementing $U(t\lambda_{\boldsymbol{c}})$, with $\lambda_{\boldsymbol{c}}$ from Eq. (robustPauliLP). However, there still remains the rest terms $R_{\boldsymbol{a},\boldsymbol{c}}$ in Eq. (73) which can be cancelled by selecting the π-pulse direction appropriately. To this end, we define the set of *robust rotation directions* of π pulses,

$$\mathcal{S}_{\boldsymbol{J}} := \Big\{ \boldsymbol{s} \in \mathbb{F}_2^n \mid \sum_{\boldsymbol{s}} (-1)^{\boldsymbol{e_a}\cdot\boldsymbol{s}} = 0,\ \forall \boldsymbol{e_a} \in \mathbb{F}_2^n \text{ with } e_{\boldsymbol{a},i} = 0 \\ \forall i \notin \mathrm{supp}(\boldsymbol{a})\ \forall \boldsymbol{a} \in \mathbb{F}_2^{2n} \text{ with } J_{\boldsymbol{a}} \neq 0 \Big\}. \tag{75}$$

This definition requires that the sum of all potential sign flips caused by the finite-pulse-time error of the nonzero interactions cancels over all rotation directions in $\mathcal{S}_{\boldsymbol{J}}$.

Proposition 7. The target time evolution $\mathrm{e}^{-\mathrm{i}tH_T}$ can be approximated by a deterministic product formula implementing $U(t\lambda_{\boldsymbol{c}})$, with $\lambda_{\boldsymbol{c}}$ from Eq. (robustPauliLP), and choosing the robust rotation directions $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{J}}$ of the π pulses. The only approximation errors are given by the ones from the Magnus approximation (Lemma 4) and the approximation error from the deterministic product formula.

The proof of Proposition 7 can be found in Appendix E.

Note that even if $\lambda_{\boldsymbol{c}} = 0$ the evolution block $U(t\lambda_{\boldsymbol{c}})$ still has to be implemented with zero free-evolution time. The number of evolution blocks can be reduced by formulating a MILP, which we explain in Sec. V E.

The robustness against finite-pulse-time errors simultaneously implies robustness against the first order effects of rotation-angle errors. This means that for a set of robust rotation directions $\boldsymbol{s} \in \mathcal{S}_{\boldsymbol{J}}$ implies the cancelation of rotation-angle errors in the first-order Taylor approximation (see Proposition 7 in Appendix E). Given a system Hamiltonian with only two-body interactions, i.e., only interactions $P_{\boldsymbol{a}}$ with $|\mathrm{supp}(\boldsymbol{a})| = 2$, a good choice for the robust rotation directions of the π pulses $\boldsymbol{s}_{(j)} \in \mathbb{F}_2^n$ for $j = 1, \ldots, \kappa$ can be found by utilizing the orthogonality property of Walsh-Hadamard matrices.

Proposition 8. Let $\kappa = 2^{\lceil \log_2(n+1) \rceil} \leq 2n$ and let $W^{(\kappa\times\kappa)}$ be the $\kappa \times \kappa$ dimensional Walsh-Hadamard matrix. Choose n distinct columns from $W^{(\kappa\times\kappa)}$ without the first

column and define the resulting partial Walsh-Hadamard matrix as $W^{(\kappa\times n)}$. Let $(-1)^{\boldsymbol{s}_{(j)}}$ be the jth row of $W^{(\kappa\times n)}$. Then, for any nonzero two-body interaction $J_{\boldsymbol{a}}\neq 0$ with $|\mathrm{supp}(\boldsymbol{a})|=2$ we have $\boldsymbol{s}_{(j)}\in\mathcal{S}_{\boldsymbol{J}}$ for all $j=1,\ldots,\kappa$.

The proof of Proposition 8 can be found in Appendix E.

E. Reducing the number of single-qubit pulses

Let the matrices $W,E\in\mathbb{R}^{r\times s}$ be either $\mathcal{W}(\boldsymbol{J})^{(r\times s)}$, $E(\boldsymbol{J})^{(r\times s)}$ or $W^{(r\times s)},E^{(r\times s)}$ as in Eq. (robustLP) or Eq. (robustPauliLP), respectively. In these LPs the sum of the finite pulse time effects for all considered evolution blocks $U(t\lambda_{\boldsymbol{c}})$ is given by the vector $E\cdot\mathbf{1}\in\mathbb{R}^r$. Moreover, the solution $\boldsymbol{\lambda}$ in Eqs. (robustLP) and (robustPauliLP) is sparse [47]. However, when mitigating the finite pulse time errors as in Theorem 5 and Proposition 7, then all the single-qubit pulse conjugations have to be implemented (with zero free-evolution time if $\lambda_{\boldsymbol{c}}=0$). The LPs can be extended with additional binary variables $\boldsymbol{z}\in\{0,1\}^s$ and an additional constraint to only consider the finite-pulse-time errors for the single-qubit pulses with nonzero free-evolution time. This yields the mixed-integer linear program (MILP) [83]

$$\begin{aligned}\text{minimize}\quad & \alpha\mathbf{1}^T\boldsymbol{\lambda}+(1-\alpha)\mathbf{1}^T\boldsymbol{z}\\ \text{subject to}\quad & W\boldsymbol{\lambda}+E\boldsymbol{z}=\boldsymbol{M},\\ & c_l\boldsymbol{z}\leq\boldsymbol{\lambda}\leq c_u\boldsymbol{z},\\ & \boldsymbol{\lambda}\in\mathbb{R}^s_{\geq 0},\quad \boldsymbol{z}\in\{0,1\}^s.\end{aligned}\tag{MILP}$$

The free parameter $\alpha\in[0,1]$ assigns weights to the minimization of the free-evolution times $\alpha=1$ or the minimization of the number of single-qubit pulse layers $\alpha=0$. $0\leq c_l<c_u$ are lower and upper bounds on the entries of $\boldsymbol{\lambda}$. The interval $[c_l,c_u]$ has to be large enough such that Eq. (MILP) has a solution.

For our robust method in Eq. (robustLP), this Eq. (MILP) reduces the number of evolution blocks in one Trotter cycle from $s\geq 2r$ to $s\approx r$. Although Eq. (MILP) is in general hard to solve, there are many powerful heuristics and software packages to solve such optimizations [53]. Moreover, the size of Eq. (MILP) can be drastically reduced with our efficient relaxation, and solving it for small instances is still feasible. Equation (MILP) is feasible only for small instances, whereas Eq. (robustLP) is efficiently solvable at the cost of more single-qubit pulse layers.

If not otherwise stated, we have used the parameters $c_l=10^{-6}$, $c_u=10^3$ and $\alpha=10^{-2}$. We have used MOSEK to solve Eq. (MILP) with parameter MSK_DPAR_MIO_TOL_REL_GAP set to 1.0 [53].

VI. DISCUSSION AND OUTLOOK

We have considered a quantum computing or quantum simulation platform that has one entangling Hamiltonian as *system Hamiltonian* and provides efficient, general, and robust methods to engineer arbitrary Hamiltonians with the same locality from it. Our methods only rely on the use of π or $\pi/2$ pulses, i.e., Pauli or single-qubit Clifford gates, and explicitly allows for robust composite pulses. They can be directly used in experiments by applying the pulse sequences generated by the provided PYTHON code [54].

The pulse sequences have been obtained by solving a suitable LP, the classical run-time of which depends polynomially only on the number of interaction terms that can be generated from the system Hamiltonian and can thus be efficiently solved for local Hamiltonians. The Pauli-conjugation method is even applicable if only partial knowledge of the system Hamiltonian is available, and can be used to cancel unwanted, but unknown, interaction terms. Moreover, the quantum simulation run-time can be reduced at the cost of a higher classical run-time, which provides a flexible trade-off.

Another major advantage of our methods is the robustness against experimental imperfections. Simulation errors introduced by finite-pulse-times can be explicitly compensated in the computation of the pulse sequence. The Clifford-conjugation method can be combined with robust composite pulses, making it robust against major experimental errors. We have discussed in detail the effect of finite-pulse-time errors and rotation angle errors and shown that these can be fully mitigated by modified pulse sequences in combination with robust composite pulses. Due to their generality and efficiency, we expect that our methods will find many applications in quantum computation and quantum simulation, such as the fast synthesis of multi-qubit gates, or analog quantum simulation for problems arising in many-body physics. Furthermore, some recent Hamiltonian learning schemes rely on "reshaping" an unknown Hamiltonian to a diagonal Hamiltonian which can be done efficiently and in a robust way with our Pauli conjugation method [37–39].

In the future, we would like to extend the efficient Hamiltonian-engineering method to fermionic systems for digital and analog quantum simulations. Another promising future research direction could be the investigation of finite-pulse-time effects for continuous robust pulses similar to Sec. V B. Moreover, the design of robust pulses tailored to the conjugation methods might be another interesting direction.

ACKNOWLEDGMENTS

We are grateful to Ivan Boldin, Patrick Huber, Markus Nünnerich, and Rodolfo Muñoz Rodriguez for a productive dialogue on the ion-trap platform, and to Matthias Zipper and Christopher Cedzich for fruitful discussions on gate designs for ion-trap platforms. Moreover, we thank Matteo Votto for a constructive exchange about his impressive work and the robust sequences therein. We also thank

Gaurav Bhole for making us aware of his results. Furthermore, we want to thank Özgün Kum for invaluable comments on our manuscript. This work has been funded by the German Federal Ministry of Research, Technology and Space (BMFTR) within the funding program "Quantum Technologies—from Basic Research to Market" via the joint project MIQRO (Grant No. 13N15522); by the QuantERA II Programme, which has received funding from the European Union (EU) H2020 research and innovation program under Grant Agreement No. 101017733 and with the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Grant No. 532779266; by the Hamburg Quantum Computing project, which is cofinanced by the EU European Regional Development Fund (ERDF) and the Fonds of the Hamburg Ministry of Science, Research, Equalities and Districts (BWFGB); and by Fujitsu Services GmbH as part of the endowed professorship "Quantum Inspired and Quantum Optimization."

DATA AVAILABILITY

The data that support the findings of this paper are openly available [54].

APPENDIX A: DETAILS FOR THE NUMERICAL SIMULATIONS

In this appendix, we provide the details for the numerical simulations in Secs. III B 1 and III B 2.

1. Simulation of a 2D lattice model

In Sec. III B 1 we have engineered a 2D lattice Hamiltonian with Ising interactions and unwanted but unknown three-body interactions. In Fig. 4, we compare the NAIVE Pauli conjugation from Sec. IV A against the ROBUST PAULI conjugation from Sec. V D. To reduce the number of required pulses we have solved Eq. (MILP) for Eq. (robustPauliLP) to obtain the relative evolution times $\boldsymbol{\lambda}$. We use multiple π-pulse rotation directions as in Proposition 8, to cancel the rest term of the finite-pulse-time error as explained in Lemma 6. Therefore, for $n = 6$ qubits $\kappa = 2^{\lceil \log_2(n+1) \rceil} = 8$ different rotation directions and thus eight evolution blocks are required within each Trotter cycle. To cancel the unwanted and unknown three-body interaction terms we apply the results from Sec. IV C to both NAIVE and ROBUST PAULI approaches.

2. Simulation of Heisenberg Hamiltonians with an ion-trap model

In Sec. III B 2 we have engineered a random Heisenberg Hamiltonian from a system Hamiltonian with Ising interactions. In Fig. 5 we compare the NAIVE Clifford conjugation from Sec. IV B with the ROBUST CLIFFORD, $\mathrm{CP}_{\mathrm{SCROFULOUS}}$, and $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ Clifford conjugations from Sec. V C. The ROBUST CLIFFORD method is only robust against the finite-pulse-time error. The $\mathrm{CP}_{\mathrm{SCROFULOUS}}$ method is additionally robust against rotation-angle errors. It is given as a combination of our robust Clifford conjugation with the SCROFULOUS composite pulses [58]. Finally, the $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ method is robust against the finite-pulse-time error, rotation angle errors, and off-resonance errors. It is given as a combination of our robust Clifford conjugation with the SCORBUTUS composite pulses [50]. We have again solved Eq. (MILP) for Eq. (robustLP) to obtain the relative evolution times $\boldsymbol{\lambda}$ and reduce the number of required pulses for the ROBUST CLIFFORD, $\mathrm{CP}_{\mathrm{SCROFULOUS}}$, and $\mathrm{CP}_{\mathrm{SCORBUTUS}}$ Clifford conjugations.

APPENDIX B: PROPERTIES OF SOLUTIONS OF EQ. (PauliLP)

In this appendix, we show the existence of solutions, provide lower and upper bounds on the relative evolution time, and discuss the tightness of these bounds. First, a direct consequence of Proposition 1 is that Eq. (PauliLP) is feasible for arbitrary Hamiltonians H_S and H_T satisfying $\mathrm{nz}(\boldsymbol{A}) \subseteq \mathrm{nz}(\boldsymbol{J})$.

Corollary B1 (Existence of solutions). Let

$$H_S = \sum_{a \in \mathbb{F}_2^{2n} \setminus \{0\}} J_a P_a \quad \text{and} \quad H_T = \sum_{a \in \mathbb{F}_2^{2n} \setminus \{0\}} A_a P_a, \tag{B1}$$

with $\mathrm{nz}(\boldsymbol{A}) \subseteq \mathrm{nz}(\boldsymbol{J})$. Then, the partial Walsh-Hadamard matrix $W^{(r \times 4^n)}$ with $r = |\mathrm{nz}(\boldsymbol{J})|$ leads to a feasible Eq. (PauliLP).

Proof. We know that all rows of the Walsh-Hadamard matrix are linearly independent. Furthermore, each row of the Walsh-Hadamard matrix sums to zero; thus $\boldsymbol{x} = \mathbf{1} > \mathbf{0}$ is a solution to $W^{(r \times 4^n)}\boldsymbol{x} = \mathbf{0}$. By Proposition 1 we know that Eq. (PauliLP) always has a feasible solution even if we consider an arbitrary subset of rows. ■

To bound the optimal solutions $\mathbf{1}^T\boldsymbol{\lambda}^*$ of Eq. (PauliLP) we define the dual LP

$$\begin{aligned} &\text{maximize} && \boldsymbol{M}^T\boldsymbol{y} \\ &\text{subject to} && (W^{(r \times 4^n)})^T\boldsymbol{y} \leq \mathbf{1}, \boldsymbol{y} \in \mathbb{R}^r. \end{aligned} \tag{B2}$$

As in the previously considered case of two-body Ising interactions [42], we have the following bounds on the value of Eq. (PauliLP).

Theorem B1 (Bounds on solutions). The optimal objective function value of Eq. (PauliLP) with a partial Walsh-Hadamard matrix $W^{(r \times 4^n)}$ is bounded by

$$||\boldsymbol{M}||_{\ell_\infty} \leq \mathbf{1}^T\boldsymbol{\lambda}^* \leq ||\boldsymbol{M}||_{\ell_1}. \tag{B3}$$

Proof. The lower bound can be verified by the fact that $W^{(r\times 4^n)}$ in Eq. (PauliLP) only has entries ± 1 and that $\boldsymbol{\lambda}^*$ is non-negative. Thus, it holds that $||\boldsymbol{M}||_{\ell_\infty} = ||W^{(r\times 4^n)}\boldsymbol{\lambda}^*|| \leq \mathbf{1}^T\boldsymbol{\lambda}^*$.

To show the upper bound, we first define the set of feasible solutions for the dual LP Eq. (B2)

$$\mathcal{F} := \left\{\boldsymbol{y} \in \mathbb{R}^r \middle| (W^{(r\times 4^n)})^T\boldsymbol{y} \leq \mathbf{1}\right\}. \tag{B4}$$

Next, consider the partial Walsh-Hadamard matrix $W^{((4^n-1)\times 4^n)}$ without the first row, corresponding to the identity Pauli term $P_{\boldsymbol{a}} = I^{\otimes n}$ with $\boldsymbol{a} = \mathbf{0}$. We show that

$$\mathcal{S} := \left\{\boldsymbol{y} \in \mathbb{R}^{4^n-1} \middle| (W^{((4^n-1)\times 4^n)})^T\boldsymbol{y} \leq \mathbf{1}\right\} \tag{B5}$$

is a simplex. A simplex is formed by affine independent vectors. Vectors of the form $(1, \boldsymbol{v}_i)^T$ are linearly dependent if and only if the vectors $\boldsymbol{v}_i$ are affine dependent. From linear dependence it follows that there is a vector $\boldsymbol{t} \neq \mathbf{0}$ such that $\sum_i t_i(1, \boldsymbol{v}_i)^T = \mathbf{0}$. Thus $\sum_i t_i = 0$ and $\sum_i \boldsymbol{v}_i = \mathbf{0}$, and the vectors $\boldsymbol{v}_i$ are affine dependent. For any n, the Walsh-Hadamard matrix $W = (\mathbf{1}, (W^{((4^n-1)\times 4^n)})^T)$ has linearly independent rows and columns. Therefore, $W^{((4^n-1)\times 4^n)}$ has affine independent columns, and $\mathcal{S}$ forms a simplex. With $\mathcal{H} = \{\boldsymbol{y} \in \mathbb{R}^{4^n-1} | ||\boldsymbol{y}||_{\ell_\infty} \leq 1\}$ we denote the $4^n - 1$ dimensional hypercube. Note that the rows of $W^{(r\times 4^n)}$ are rows of the $4^n \times 4^n$ dimensional Walsh-Hadamard matrix. We define the embedding $g : \mathbb{R}^r \to \mathbb{R}^{(4^n)-1}$ by appending zeros, such that $(W^{(r\times 4^n)})^T\boldsymbol{y} = (W^{((4^n-1)\times 4^n)})^T g(\boldsymbol{y})$. Therefore the objective value of the dual LP Eq. (B2) can be upper bounded as follows:

$$\begin{aligned}\max_{\boldsymbol{y}\in\mathcal{F}}\langle \boldsymbol{M}\boldsymbol{y}\rangle &= \max_{g(\boldsymbol{y})\in\mathcal{S}}\langle g(\boldsymbol{M})g(\boldsymbol{y})\rangle \\ &\leq \max_{\boldsymbol{x}\in\mathcal{S}}\langle g(\boldsymbol{M})\boldsymbol{x}\rangle \\ &\leq \max_{\boldsymbol{x}\in\mathcal{H}}\langle g(\boldsymbol{M})\boldsymbol{x}\rangle = ||\boldsymbol{M}||_{\ell_1}.\end{aligned} \tag{B6}$$

The upper bound for $\mathbf{1}^T\boldsymbol{\lambda}^*$ follows by strong duality. ■

Next, we show a (not complete) set of instances of Eq. (PauliLP), yielding solutions $\boldsymbol{\lambda}^*$ which satisfy the upper bound of Theorem B1. Such $\boldsymbol{M}$ constitute the worst cases.

Proposition B1. If $\boldsymbol{M} \in \left\{-\boldsymbol{w}_i \middle| \boldsymbol{w}_i \text{ is the } i\text{th column of } W^{((4^n-1)\times 4^n)}\right\}$, then $\mathbf{1}^T\boldsymbol{\lambda}^* = ||\boldsymbol{M}||_{\ell_1}$.

Proof. Let $\boldsymbol{y} = -\boldsymbol{w}_i$ be a negative column of $W^{((4^n-1)\times 4^n)}$. Then, by orthogonality of the rows or columns of the Walsh-Hadamard matrix, we obtain

$$\left((W^{((4^n-1)\times 4^n)})^T(-\boldsymbol{w}_i)\right)_j = \begin{cases}-(4^n-1), & i = j,\\ 1, & \text{otherwise.}\end{cases} \tag{B7}$$

Therefore, $\boldsymbol{y} = -\boldsymbol{w}_i$ is a feasible solution $(W^{((4^n-1)\times 4^n)})^T\boldsymbol{y} \leq \mathbf{1}$ to the dual LP Eq. (B2). The dual-objective-function value is $\boldsymbol{M}^T\boldsymbol{y} = ||\boldsymbol{w}_i||_{\ell_1} = 4^n - 1$. Next, we show a feasible solution of Eq. (PauliLP). For a $\boldsymbol{M} = -\boldsymbol{w}_i$, define

$$\lambda_j = \begin{cases}0, & i = j,\\ 1, & \text{otherwise.}\end{cases} \tag{B8}$$

Clearly, this satisfies $W^{((4^n-1)\times 4^n)}\boldsymbol{\lambda} = -\boldsymbol{w}_i$. Furthermore, the primal-objective-function value $\mathbf{1}^T\boldsymbol{\lambda} = 4^n - 1$ is the same as for the dual-objective-function value, which shows optimality. It is easy to see that $||\boldsymbol{M}||_{\ell_1} = 4^n - 1$. ■

APPENDIX C: COMMENTS ON THE EFFICIENT RELAXATION FOR THE PAULI CONJUGATION

As mentioned in the main text, Wendel's theorem is applicable to spherical symmetric distributions. This would imply that sampling a certain column from $W^{(r\times 4^n)}$ has the same probability as sampling the same column with a flipped sign. Recall that for a partial Walsh-Hadamard matrix $W^{(r\times 4^n)}_{\boldsymbol{ab}} = (-1)^{\langle \boldsymbol{a},\boldsymbol{b}\rangle}$, with $\langle \boldsymbol{a},\boldsymbol{b}\rangle = \boldsymbol{a}_x \cdot \boldsymbol{b}_z + \boldsymbol{a}_z \cdot \boldsymbol{b}_x$, it holds that

$$W^{(r\times 4^n)}_{\boldsymbol{ab}} = -W^{(r\times 4^n)}_{\boldsymbol{a}\bar{\boldsymbol{b}}} \quad \forall \boldsymbol{a},\boldsymbol{b} \in \mathbb{F}_2^{2n}, \tag{C1}$$

for $|a| = 1 \bmod 2$ and with the binary complement $\bar{\boldsymbol{b}}$ given element-wise given by $\bar{x} := 1 - x$ for any $x \in \{0, 1\}$. If the decomposition of H_S has only terms $J_{\boldsymbol{a}}P_{\boldsymbol{a}}$ with odd $|a|$, then we can apply Wendel's theorem directly. In this case, we have a success probability of $1/2$ of finding a feasible $W^{(r\times 2r)}$ (with r interactions) if we sample $2r$ many $\boldsymbol{b} \in \mathbb{F}_2^{2n}$ uniformly random. However, an example of such a Hamiltonian with two-body interactions is

$$H = \sum_{i=1}^n \left(J_i^X X_i + J_i^Y Y_i\right) + \sum_{i\neq j}^n \left(J_{ij}^{XZ} X_i Z_j + J_{ij}^{YZ} Y_i Z_j\right) \tag{C2}$$

and with commuting interactions is

$$H = \sum_{i=1}^n J_i^X X_i + \sum_{ijk}^n J_{ijk}^{XXX} X_i X_j X_k, \tag{C3}$$

with arbitrary coupling constants. Unfortunately, for a general Hamiltonian, we cannot use Wendel's theorem to construct a relaxation for Eq. (PauliLP).

Recently, lower bounds on $p_{s,\boldsymbol{x}}$ have been proposed for arbitrary distributions [62]. However, their results rely on the half-space depth (or Tukey depth), which is hard to compute. It is a measure of how extreme a point is with respect to a distribution of random points.

Definition C1 (Half-space depth). Let $\boldsymbol{x}$ be an arbitrary r-dimensional random vector. Then, the half-space depth at the origin is defined as

$$\alpha_{\boldsymbol{x}} := \inf_{\|\boldsymbol{c}\|=1} \mathbb{P}(\boldsymbol{c}^T\boldsymbol{x} \leq \boldsymbol{0}). \tag{C4}$$

The half-space depth $\alpha_{\boldsymbol{x}}$ is the minimum (fraction) number of points in a half-space with the origin on the boundary.

Theorem C1 (Theorem 14 of [62]). Let $\boldsymbol{x}$ be an arbitrary r-dimensional random vector. Then, for each positive integer $s \geq 3r/\alpha_{\boldsymbol{x}}$, we have

$$p_{s,\boldsymbol{x}} > 1 - \frac{1}{2^r}. \tag{C5}$$

Let $\boldsymbol{w}$ be a r-dimensional random vector drawn uniformly from $\mathrm{col}(W^{(r\times 4^n)})$. From Corollary B1, we find the trivial lower bound $1/4^n \leq \alpha_{\boldsymbol{w}}$, since at least one point is in an arbitrary half-space with the origin on its boundary. Finding a constant lower bound $1/\beta \leq \alpha_{\boldsymbol{w}}$ would imply that for $s \geq 3r\beta$ we find a feasible $W^{(r\times s)}$ with high probability. One needs to show that at least $4^{\mathrm{O}(n)}$ points are contained in an arbitrary half-space with the origin on its boundary.

APPENDIX D: PROOFS FOR THE GENERAL ROBUST CONJUGATION METHOD

Here, we provide the proofs for Sec. V B. Some proofs provide more technical details with the aim of easy implementation into a programming language. For the sake of easy readability we repeat the definitions from the main text.

Definition 3. Let $\mathcal{G} \subset \mathrm{Herm}(\mathbb{C}^2)$ be a set of *single-qubit pulse generators*. A *single-qubit pulse layer* is given by a *local generator* h_i chosen from $\mathcal{G}$ for each qubit $i \in [n]$, the *(single-qubit) rotation directions* $\boldsymbol{s} \in \mathbb{F}_2^n$ and the *finite pulse time* $t_p > 0$. The Hamiltonian generating the single-qubit pulse layer on n qubits is given by

$$H(t_p, \boldsymbol{s}, \boldsymbol{h}) := \frac{1}{t_p}\sum_{i=1}^{n}(-1)^{s_i}h_i, \tag{61}$$

with $\boldsymbol{h} := (h_1, \ldots, h_n)$. The single-qubit pulse layer Hamiltonian is completely specified by the tuple $c := (t_p, \boldsymbol{s}, \boldsymbol{h})$, and we write $H_c := H(t_p, \boldsymbol{s}, \boldsymbol{h})$. This generates the evolution operator for one single-qubit pulse layer

$$\boldsymbol{S}_c(t) := \mathrm{e}^{-\mathrm{i}tH_c}, \quad \text{and} \quad \boldsymbol{S}_c := \boldsymbol{S}_c(t_p). \tag{62}$$

More generally, we consider a sequence of n_L single-qubit pulse layers and introduce the layer index $\ell \in [n_L]$. The evolution for the ℓth single-qubit pulse layer is specified by the tuple $c^{(\ell)} := (t_p^{(\ell)}, \boldsymbol{s}^{(\ell)}, \boldsymbol{h}^{(\ell)})$, and we define the tuple $\boldsymbol{c} := (c^{(1)}, \ldots, c^{(n_L)})$. Moreover, we define the *single-qubit pulse block* and the *partial single-qubit pulse block* as

$$\boldsymbol{S}_{\boldsymbol{c}} := \overrightarrow{\prod_{\ell=1}^{n_L}} \boldsymbol{S}_{c^{(\ell)}} \tag{63}$$

and

$$\boldsymbol{S}_{\boldsymbol{c}^{\geq D}} := \begin{cases} \overrightarrow{\prod}_{\ell=D}^{n_L} \boldsymbol{S}_{c^{(\ell)}}, & 1 \leq D \leq n_L, \\ \mathbb{1}, & \text{otherwise}, \end{cases} \tag{64}$$

respectively. Similarly, we define the pulse time for the single-qubit pulse block $T_p := \sum_{\ell=1}^{n_L} t_p^{(\ell)}$ and for the partial single-qubit pulse block $T_p^{\leq D} := \sum_{\ell=1}^{D} t_p^{(\ell)}$.

We consider the time evolution

$$U(t\lambda_{\boldsymbol{c}}) = \left(\overrightarrow{\prod_{\ell=1}^{n_L}} \mathrm{e}^{-\mathrm{i}t_p^{(\ell)}(H_S - H_{c^{(\ell)}})}\right) \times \mathrm{e}^{-\mathrm{i}t\lambda_{\boldsymbol{c}}H_S} \left(\overrightarrow{\prod_{\ell=1}^{n_L}} \mathrm{e}^{-\mathrm{i}t_p^{(\ell)}(H_S + H_{c^{(\ell)}})}\right), \tag{D1}$$

with the finite duration of the ℓth pulse layer $0 < t_p^{(\ell)}$.

Lemma 4. The approximation of $U(t\lambda_{\boldsymbol{c}})$ in first-order Magnus expansion is given by $\mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)}$ with

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\lambda_{\boldsymbol{c}}\boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}, \tag{68}$$

where

$$H_{\mathrm{err},\boldsymbol{c}} := 2\sum_{\ell=1}^{n_L} \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^\dagger \int_0^{t_p^{(\ell)}} \boldsymbol{S}_{c^{(\ell)}}^\dagger(t) H_S \boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t\, \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}. \tag{69}$$

The average Hamiltonian $H_{\mathrm{av},\boldsymbol{c}}(t)$ has the same locality as H_S. Moreover, the error for truncating the Magnus expansion after the first order can be bounded in the spectral norm by

$$\|U(t\lambda_{\boldsymbol{c}}) - \mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)}\| \leq \mathrm{O}((2T_p + t\lambda_{\boldsymbol{c}})^2\|H_S\|^2). \tag{70}$$

Proof. We consider the time evolution during the ℓth layer of single-qubit pulses in the interaction frame with respect to the single-qubit pulse Hamiltonian $H_{c^{(\ell)}}$ [66,81]

$$\mathrm{e}^{-\mathrm{i}t_p^{(\ell)}(H_S\pm H_{c^{(\ell)}})} = \mathrm{e}^{\mp\mathrm{i}t_p^{(\ell)}H_{c^{(\ell)}}}\mathcal{U}_{\pm}^{(\ell)}(t_p^{(\ell)}) = \boldsymbol{S}_{c^{(\ell)}}^{\pm1}\mathcal{U}_{\pm}^{(\ell)}(t_p^{(\ell)}), \tag{D2}$$

with the interaction-frame propagator $\mathcal{U}_{\pm}^{(\ell)}(t_p^{(\ell)})$. The interaction-frame propagator has to fulfill

$$\frac{d\mathcal{U}_{\pm}^{(\ell)}(t)}{dt} = -\mathrm{i}\left(\boldsymbol{S}_{c^{(\ell)}}^{\mp1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}^{\pm1}(t)\right)\mathcal{U}_{\pm}^{(\ell)}(t). \tag{D3}$$

Inserting Eq. (D2) into Eq. (D1) yields

$$U(t\boldsymbol{\lambda_c}) = \left(\prod_{\ell=1}^{\overrightarrow{n_L}}\boldsymbol{S}_{c^{(\ell)}}^{-1}\mathcal{U}_{-}^{(\ell)}(t_p^{(\ell)})\right) \times \mathrm{e}^{-\mathrm{i}t\boldsymbol{\lambda_c}H_S}\left(\prod_{\ell=1}^{\overrightarrow{n_L}}\boldsymbol{S}_{c^{(\ell)}}\mathcal{U}_{+}^{(\ell)}(t_p^{(\ell)})\right). \tag{D4}$$

We define $\tilde{\mathcal{U}}_{-}^{(\ell)}(t)$ as

$$\frac{d\tilde{\mathcal{U}}_{-}^{(\ell)}(t)}{dt} := \boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}^{-1}\frac{d\mathcal{U}(\ell)_{-}(t)}{dt}\boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}} = -\mathrm{i}\left(\boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}^{-1}\boldsymbol{S}_{c^{(\ell)}}(t)H_S\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)\boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}\right)\tilde{\mathcal{U}}_{-}^{(\ell)}(t) \tag{D5}$$

and $\tilde{\mathcal{U}}_{+}^{(\ell)}(t)$ as

$$\frac{d\tilde{\mathcal{U}}_{+}^{(\ell)}(t)}{dt} := \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\frac{d\mathcal{U}(\ell)_{+}(t)}{dt}\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}} = -\mathrm{i}\left(\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}\right)\tilde{\mathcal{U}}_{+}^{(\ell)}(t), \tag{D6}$$

where $\tilde{\mathcal{U}}_{-}^{(\ell)}(t)$ and $\tilde{\mathcal{U}}_{+}^{(\ell)}(t)$ are defined such that we can commute the exact evolution of the single-qubit pulse layers to the free evolution under H_S. It directly follows that

$$U(t\boldsymbol{\lambda_c}) = \left(\prod_{\ell=1}^{\overrightarrow{n_L}}\tilde{\mathcal{U}}_{-}^{(\ell)}(t)\right)\mathrm{e}^{-\mathrm{i}t\boldsymbol{\lambda_c}\boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}}}\left(\prod_{\ell=1}^{\overrightarrow{n_L}}\tilde{\mathcal{U}}_{+}^{(\ell)}(t)\right). \tag{D7}$$

The Hamiltonian governing the evolution of $U(t\boldsymbol{\lambda_c})$ is defined piecewise as

$$H_{\boldsymbol{c}}(\tilde{t}) := \begin{cases} \boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\boldsymbol{S}_{c^{(\ell)}}^{-1}(\tilde{t})H_S\boldsymbol{S}_{c^{(\ell)}}(\tilde{t})\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}, & T_p^{\leq(\ell-1)} \leq \tilde{t} < T_p^{\leq\ell}, \\ \boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}}, & T_p \leq \tilde{t} < T_p + t\boldsymbol{\lambda_c}, \\ \boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}^{-1}\boldsymbol{S}_{c^{(\ell)}}(\tilde{t})H_S\boldsymbol{S}_{c^{(\ell)}}^{-1}(\tilde{t})\boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}, & T_p + T_p^{\leq(\ell-1)} + t\boldsymbol{\lambda_c} \leq \tilde{t} < T_p + T_p^{\leq\ell} + t\boldsymbol{\lambda_c}. \end{cases} \tag{D8}$$

The effective Hamiltonian is approximated by the first-order term in the Magnus expansion up to time $T = 2T_p + t\boldsymbol{\lambda_c}$, yielding the time-independent average Hamiltonian

$$\begin{aligned} H_{\mathrm{av},\boldsymbol{c}}(t) &= \int_0^T H_{\boldsymbol{c}}(\tilde{t})\mathrm{d}\tilde{t} = t\boldsymbol{\lambda_c}\boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}} + \sum_{\ell=1}^{n_L}\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\int_0^{t_p^{(\ell)}}\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}} + \boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}}^{-1}\int_0^{t_p^{(\ell)}}\boldsymbol{S}_{c^{(\ell)}}(t)H_S\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)\mathrm{d}t\boldsymbol{S}_{\boldsymbol{c}^{\geq\ell}} \\ &= t\boldsymbol{\lambda_c}\boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}} + 2\sum_{\ell=1}^{n_L}\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\int_0^{t_p^{(\ell)}}\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}. \end{aligned} \tag{D9}$$

We denote the term corresponding to the finite-pulse-time error by

$$H_{\mathrm{err},\boldsymbol{c}} := 2\sum_{\ell=1}^{n_L}\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}^{-1}\int_0^{t_p^{(\ell)}}\boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t\,\boldsymbol{S}_{\boldsymbol{c}^{\geq(\ell+1)}}. \tag{D10}$$

Then, the average Hamiltonian is

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\boldsymbol{\lambda_c}\boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}. \tag{D11}$$

Note that conjugation of the system Hamiltonian H_S with single-qubit operations as in $\boldsymbol{S}_{\boldsymbol{c}}^{-1} H_S \boldsymbol{S}_{\boldsymbol{c}}$ and $H_{\mathrm{err},\boldsymbol{c}}$ always preserves the locality of H_S.

Finally, we prove the error bounds on the truncation of the Magnus expansion after the first order. The kth-order term of the Magnus expansion can be written as

$$H_{\mathrm{av}}^{(k)} = \sum_{\sigma \in S_k} (-1)^{d_b} \frac{d_a! d_b!}{k!} \int_0^T \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \ldots \times \int_0^{\tau_{k-1}} \mathrm{d}\tau_k H_{\boldsymbol{c}}(\tau_{\sigma(1)}) H_{\boldsymbol{c}}(\tau_{\sigma(2)}) \ldots H_{\boldsymbol{c}}(\tau_{\sigma(k)}), \tag{D12}$$

where $T = 2T_p + t\lambda_{\boldsymbol{c}}$, S_k denotes the group of permutations σ of the set $[k]$, the number of ascents d_a and descents d_b are defined as

$$d_a := |\{i \in [k-1] | \sigma(i) < \sigma(i+1)\}| \quad \text{and} \quad d_b := |\{i \in [k-1] | \sigma(i) > \sigma(i+1)\}|, \tag{D13}$$

and it holds that $d_a + d_b = k$ [84]. Finally, we argue that

$$\|H_{\mathrm{av}}^{(k)}\| \leq \mathrm{O}\left((2T_p + t\lambda_{\boldsymbol{c}})^k \| \max_{\tau \in [0, 2T_p + t\lambda_{\boldsymbol{c}}]} H_{\boldsymbol{c}}(\tau) \|^k \right) = \mathrm{O}((2T_p + t\lambda_{\boldsymbol{c}})^k \|H_S\|^k), \tag{D14}$$

where the equality follows from the definition of $H_{\boldsymbol{c}}(\tau) = U^{-1} H_S U$ for some unitaries U and the invariance of the spectral norm under unitary transformation. The error bound in the spectral norm follows from the Duhamel principle in Ref. [82, App. A]:

$$\| U(t\lambda_{\boldsymbol{c}}) - \mathrm{e}^{-\mathrm{i} H_{\mathrm{av},\boldsymbol{c}}(t)} \| \leq \sum_{k=2}^{\infty} \|H_{\mathrm{av}}^{(k)}\| = \mathrm{O}((2T_p + t\lambda_{\boldsymbol{c}})^2 \|H_S\|^2), \tag{D15}$$

assuming that $\|H_{\mathrm{av}}^{(2)}\| > \sum_{k=3}^{\infty} \|H_{\mathrm{av}}^{(k)}\|$, i.e., the Magnus expansion converges. ■

Lemma D1. Let $\mathcal{W}(\boldsymbol{J})^{(r\times s)}, E(\boldsymbol{J})^{(r\times s)} \in \mathbb{R}^{r\times s}$ be the matrices representing the ideal conjugation and the finite-pulse-time error in the Pauli basis, respectively. The entries are given by

$$\mathcal{W}(\boldsymbol{J})_{\boldsymbol{ac}}^{(r\times s)} := \frac{1}{2^n} \mathrm{Tr}\left(P_{\boldsymbol{a}} \left(\boldsymbol{S}_{\boldsymbol{c}}^{\dagger} H_S \boldsymbol{S}_{\boldsymbol{c}}\right)\right) \quad \text{and} \quad E(\boldsymbol{J})_{\boldsymbol{ac}}^{(r\times s)} := \frac{1}{2^n} \mathrm{Tr}\left(P_{\boldsymbol{a}} H_{\mathrm{err},\boldsymbol{c}}\right) \tag{D16}$$

and can be calculated in polynomial time in the number of qubits for a local system Hamiltonian H_S.

Proof. Recall that the system Hamiltonian has the form

$$H_S = \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} J_{\boldsymbol{a}} P_{\boldsymbol{a}}. \tag{D17}$$

Let $H_{c^{(\ell)}} = \theta^{(\ell)}/t_p^{(\ell)} \sum_{i=1}^n (-1)^{s_i^{(\ell)}} h_i^{(\ell)}$ be an arbitrary layer of single-qubit rotations, with rotation angle $\theta^{(\ell)}$, pulse duration $t_p^{(\ell)}$, generators $h_i^{(\ell)} := p_{x,i}^{(\ell)} X + p_{y,i}^{(\ell)} Y + p_{z,i}^{(\ell)} Z$, and rotation direction $s_i^{(\ell)} \in \mathbb{F}_2$. For the sake of clear notation we omit the layer index (ℓ) for now. The single-qubit pulse on the ith qubit can be written as

$$S_{c_i}(t) := \mathrm{e}^{-\mathrm{i} t \frac{\theta}{t_p} h_i} = \cos\left(t \frac{\theta}{t_p}\right) I - \mathrm{i}(-1)^{s_i} \sin\left(t \frac{\theta}{t_p}\right) h_i. \tag{D18}$$

Then, $\boldsymbol{S}_c(t) = \mathrm{e}^{-\mathrm{i} t H_c} = \bigotimes_{i=1}^n S_{c_i}(t)$. The conjugation with a single-qubit pulse layer changes the interaction term as

$$\boldsymbol{S}_c^{-1}(t) J_{\boldsymbol{a}} P_{\boldsymbol{a}} \boldsymbol{S}_c(t) = J_{\boldsymbol{a}} \bigotimes_{i=1}^n S_{c_i}^{-1}(t) P_{\boldsymbol{a}_i} S_{c_i}(t). \tag{D19}$$

The effect of such a conjugation on the ith qubit is given by

$$S_{c_i}^{-1}(t) P_{\boldsymbol{a}_i} S_{c_i}(t) = \cos^2(\tilde{\theta}) P_{\boldsymbol{a}_i} + \sin^2(\tilde{\theta}) h_i P_{\boldsymbol{a}_i} h_i + \mathrm{i}[h_i, P_{\boldsymbol{a}_i}](-1)^{s_i} \sin(\tilde{\theta}) \cos(\tilde{\theta}), \tag{D20}$$

with $\tilde{\theta} := t(\theta/t_p)$. Equation (D20) can be further decomposed into Pauli terms $S_{c_i}^{-1}(t) P_{\boldsymbol{a}_i} S_{c_i}(t) = g_{x,i}(\tilde{\theta}) X + g_{y,i}(\tilde{\theta}) Y + g_{z,i}(\tilde{\theta}) Z$, with

$$g_{x,i}(\tilde{\theta}) := \begin{cases} \cos^2(\tilde{\theta}) + (p_{x,i}^2 - p_{y,i}^2 - p_{z,i}^2) \sin^2(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = X, \\ 2p_{x,i}p_{y,i} \sin^2(\tilde{\theta}) + 2p_{z,i}(-1)^{s_i} \sin(\tilde{\theta}) \cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Y, \\ 2p_{x,i}p_{z,i} \sin^2(\tilde{\theta}) - 2p_{y,i}(-1)^{s_i} \sin(\tilde{\theta}) \cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Z, \end{cases}$$

$$g_{y,i}(\tilde{\theta}) := \begin{cases} 2p_{x,i}p_{y,i} \sin^2(\tilde{\theta}) + 2p_{z,i}(-1)^{s_i} \sin(\tilde{\theta}) \cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = X, \\ \cos^2(\tilde{\theta}) + (p_{y,i}^2 - p_{x,i}^2 - p_{z,i}^2) \sin^2(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Y, \\ 2p_{y,i}p_{z,i} \sin^2(\tilde{\theta}) - 2p_{x,i}(-1)^{s_i} \sin(\tilde{\theta}) \cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Z, \end{cases} \tag{D21}$$

$$g_{z,i}(\tilde{\theta}) := \begin{cases} 2p_{x,i}p_{z,i}\sin^2(\tilde{\theta}) - 2p_{y,i}(-1)^{s_i}\sin(\tilde{\theta})\cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = X, \\ 2p_{y,i}p_{z,i}\sin^2(\tilde{\theta}) + 2p_{x,i}(-1)^{s_i}\sin(\tilde{\theta})\cos(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Y, \\ \cos^2(\tilde{\theta}) + (p_{z,i}^2 - p_{x,i}^2 - p_{y,i}^2)\sin^2(\tilde{\theta}), & \text{if } P_{\boldsymbol{a}_i} = Z, \end{cases}$$

and if $P_{\boldsymbol{a}_i} = I$, then $S_{c_i}^{-1}(t)P_{\boldsymbol{a}_i}S_{c_i}(t) = I$. Assume that the interaction $J_{\boldsymbol{a}}P_{\boldsymbol{a}}$ is k-local, i.e., the interaction acts on the qubits $i \in \mathrm{supp}(\boldsymbol{a})$ and $|\mathrm{supp}(\boldsymbol{a})| = k$. Applying $A \otimes (B + C) = A \otimes B + A \otimes C$ the conjugation in Eq. (D19) yields

$$\bigotimes_{i=1}^{n} S_{c_i}^{-1}(t)P_{\boldsymbol{a}_i}S_{c_i}(t) = \sum_{\substack{\tilde{\boldsymbol{a}}\in\mathbb{F}_2^n \\ \mathrm{supp}(\tilde{\boldsymbol{a}})=\mathrm{supp}(\boldsymbol{a})}} \left(\prod_{i\in\tilde{\boldsymbol{a}}} g_{\tilde{\boldsymbol{a}}_i,i}(\tilde{\theta})\right) P_{\tilde{\boldsymbol{a}}}, \tag{D22}$$

where we identify $g_{(1,0),i} = g_{x,i}$, $g_{(1,1),i} = g_{y,i}$, and $g_{(0,1),i} = g_{z,i}$. This sum has at most 3^k terms, which is constant for a system Hamiltonian H_S with a fixed locality.

With that we are ready to compute $E(\boldsymbol{J})_{\boldsymbol{ac}}^{(r\times s)}$ by calculating the Pauli coefficients of

$$H_{\mathrm{err},\boldsymbol{c}} = 2\sum_{\ell=1}^{S} \boldsymbol{S}_{\boldsymbol{c}\geq(\ell+1)}^{-1} \int_0^{t_p^{(\ell)}} \boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t\, \boldsymbol{S}_{\boldsymbol{c}\geq(\ell+1)}. \tag{D23}$$

We start with the integral

$$\begin{aligned} &\int_0^{t_p^{(\ell)}} \boldsymbol{S}_{c^{(\ell)}}^{-1}(t)H_S\boldsymbol{S}_{c^{(\ell)}}(t)\mathrm{d}t \\ &= \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} J_{\boldsymbol{a}} \int_0^{t_p^{(\ell)}} \bigotimes_{i=1}^{n} S_{c_i^{(\ell)}}^{-1}(t)P_{\boldsymbol{a}_i}S_{c_i^{(\ell)}}(t)\mathrm{d}t \\ &= \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} \sum_{\substack{\tilde{\boldsymbol{a}}\in\mathbb{F}_2^n \\ \mathrm{supp}(\tilde{\boldsymbol{a}})=\mathrm{supp}(\boldsymbol{a})}} J_{\boldsymbol{a}} \frac{t_p^{(\ell)}}{\theta^{(\ell)}} \int_0^{\theta^{(\ell)}} \left(\prod_{i\in\tilde{\boldsymbol{a}}} g_{\tilde{\boldsymbol{a}}_i,i}(\tilde{\theta})\right) \mathrm{d}\tilde{\theta} P_{\tilde{\boldsymbol{a}}} \\ &=: \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} E_{\boldsymbol{a}}^{(\ell)} P_{\boldsymbol{a}}, \end{aligned} \tag{D24}$$

where $E_{\boldsymbol{a}}^{(\ell)}$ can be efficiently calculated by integrating trigonometric polynoms and summing all contributions from the terms with the same locality $\mathrm{supp}(\tilde{\boldsymbol{a}}) = \mathrm{supp}(\boldsymbol{a})$. Next, the conjugation $\boldsymbol{S}_{\boldsymbol{c}\geq(\ell+1)}^{-1}(\cdot)\boldsymbol{S}_{\boldsymbol{c}\geq(\ell+1)}$ in Eq. (D23) corresponds to applying Eq. (D22) for each $\tilde{\ell} = 1, \ldots, \ell+1$ to $E_{\boldsymbol{a}}^{(\ell)}P_{\boldsymbol{a}}$ with $t = t_p^{(\tilde{\ell})}$ or $\tilde{\theta} = \theta^{(\tilde{\ell})}$. Then, we obtain

$$H_{\mathrm{err},\boldsymbol{c}} = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} E(\boldsymbol{J})_{\boldsymbol{ac}}^{(r\times s)} P_{\boldsymbol{a}}. \tag{D25}$$

The entries of $\mathcal{W}(\boldsymbol{J})^{(r\times s)}$ can be calculated similarly by applying Eq. (D22) for each $\ell = 1, \ldots, n_L$ to $J_{\boldsymbol{a}}P_{\boldsymbol{a}}$ with $t = t_p^{(\ell)}$ or $\tilde{\theta} = \theta^{(\ell)}$. Together, we obtain

$$\boldsymbol{S}_{\boldsymbol{c}}^{-1}H_S\boldsymbol{S}_{\boldsymbol{c}} = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} \mathcal{W}(\boldsymbol{J})_{\boldsymbol{ac}}^{(r\times s)} P_{\boldsymbol{a}}. \tag{D26}$$

■

Theorem 5. The target time evolution $\mathrm{e}^{-\mathrm{i}tH_T}$ can be efficiently approximated by a deterministic product formula implementing a product of $U(t\lambda_{\boldsymbol{c}})$, with $\lambda_{\boldsymbol{c}}$ and the corresponding pulse block $\boldsymbol{S}_{\boldsymbol{c}}$ from Eq. (robustLP). Moreover, this approximation is robust against the finite-pulse-time effect. The only approximation errors are given by the ones from the Magnus approximation (Lemma 4) and the approximation error from the deterministic product formula.

Proof. The chosen product formula determines the number of implemented single-qubit conjugations $U(t\lambda_{\boldsymbol{c}})$ for each $\boldsymbol{c}$. Let this number be $n_{\boldsymbol{c}}$. The finite-pulse-time error term in Eq. (robustLP) has to be rescaled to account for $n_{\boldsymbol{c}}$ since each implementation of $U(t\lambda_{\boldsymbol{c}})$ causes the associated finite pulse time error. This can be done by multiplying the $\boldsymbol{c}$th column of $E(\boldsymbol{J})^{(r\times s)}$ by $n_{\boldsymbol{c}}$.

For simplicity, we assume the first-order Trotter approximation in Eq. (18). To account for the number of implemented single-qubit conjugations we have to rescale $E(\boldsymbol{J})^{(r\times s)}$ by $n_{\boldsymbol{c}} = n_{\mathrm{Tro}}$ for all $\boldsymbol{c}$. Let $tH_{\mathrm{av}} := \sum_{\boldsymbol{c}}^{s} H_{\mathrm{av},\boldsymbol{c}}(t)$, with $\lambda_{\boldsymbol{c}}$ from Eq. (robustLP). Let the target Hamiltonian be

$$H_T = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} A_{\boldsymbol{a}}P_{\boldsymbol{a}}. \tag{D27}$$

By the constraint of Eq. (robustLP), we have $H_{\mathrm{av}} = H_T$. The time evolution governed by the target Hamiltonian H_T can be approximated with

$$\begin{aligned} \mathrm{e}^{-\mathrm{i}tH_T} = \mathrm{e}^{-\mathrm{i}tH_{\mathrm{av}}} &\approx \left(\overrightarrow{\prod_{\boldsymbol{c}}} \mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}\left(\frac{t}{n_{\mathrm{Tro}}}\right)}\right)^{n_{\mathrm{Tro}}} \\ &\approx \left(\overrightarrow{\prod_{\boldsymbol{c}}} U(t\lambda_{\boldsymbol{c}}/n_{\mathrm{Tro}})\right)^{n_{\mathrm{Tro}}}, \end{aligned} \tag{D28}$$

where the first approximation is given by the Trotter scheme and the second approximation is given by the first-order Magnus expansion from Lemma 4. ■

APPENDIX E: PROOFS FOR THE ROBUST PAULI CONJUGATION METHOD

Lemma 6. Consider all labels $\boldsymbol{a} \in \mathbb{F}_2^{2n}$ for the nonzero interactions $J_{\boldsymbol{a}} \neq 0$ in the system Hamiltonian. Then, the approximation of $U(t\lambda_{\boldsymbol{c}})$ for the Pauli conjugation in the first-order Magnus expansion is given by $\mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}}(t)}$ with

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\lambda_{\boldsymbol{c}}\boldsymbol{S}_{\boldsymbol{c}}^{\dagger}H_S\boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}, \tag{72}$$

where

$$H_{\mathrm{err},\boldsymbol{c}} = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} \left(J_{\boldsymbol{a}}E_{\boldsymbol{a},\boldsymbol{c}}^{(r\times s)}P_{\boldsymbol{a}} + R_{\boldsymbol{a},\boldsymbol{c}}\right). \tag{73}$$

We call the first term in Eq. (73) the *interaction term*, with

$$E_{\boldsymbol{a},\boldsymbol{c}}^{(r\times s)} := \frac{4t_p}{\pi}\int_0^{\frac{\pi}{2}} \left(\prod_{i\in\mathrm{supp}(\boldsymbol{a})} (\cos^2(\theta) + (-1)^{\langle \boldsymbol{a}_i,\boldsymbol{b}_i\rangle}\sin^2(\theta))\right)\mathrm{d}\theta, \tag{74}$$

which we collect as entries of the matrix $E^{(r\times s)} \in \mathbb{R}^{r\times s}$. We call the second term $R_{\boldsymbol{a},\boldsymbol{c}}$ the *rest term*, and it is proportional to $(-1)^{\boldsymbol{e}_{\boldsymbol{a}}\cdot\boldsymbol{s}}$, with $\boldsymbol{s} \in \mathbb{F}_2^n$ representing the chosen rotation direction of the π pulses and $\boldsymbol{e}_{\boldsymbol{a}} \in \mathbb{F}_2^n$ encodes the sign flips due to the finite-pulse-time error such that $e_{\boldsymbol{a},i} = 0$ for all $i \notin \mathrm{supp}(\boldsymbol{a})$.

Proof. From Lemma 4 with $n_L = 1$ and $H_{\boldsymbol{c}} = H_{c^{(1)}}$ we obtain the Hamiltonian corresponding to the finite-pulse-time effect,

$$H_{\mathrm{err},\boldsymbol{c}} := 2\int_0^{t_p} \mathrm{e}^{\mathrm{i}tH_{\boldsymbol{c}}}H_S\mathrm{e}^{-\mathrm{i}tH_{\boldsymbol{c}}}\mathrm{d}t. \tag{E1}$$

Recall that the system Hamiltonian has the form

$$H_S = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} J_{\boldsymbol{a}}P_{\boldsymbol{a}}. \tag{E2}$$

Before we compute $H_{\mathrm{err},\boldsymbol{c}}$ we investigate the conjugation $\mathrm{e}^{\mathrm{i}tH_{\boldsymbol{c}}}H_S\mathrm{e}^{-\mathrm{i}tH_{\boldsymbol{c}}}$ for a single-qubit. A π pulse on the ith qubit can be written as

$$\begin{aligned} S_{c_i}(t) &:= \mathrm{e}^{-\mathrm{i}t\frac{\pi}{2t_p}(-1)^{s_i}P_{\boldsymbol{b}_i}} \\ &= \cos\left(\frac{\pi}{2}\frac{t}{t_p}\right)I - \mathrm{i}(-1)^{s_i}P_{\boldsymbol{b}_i}\sin\left(\frac{\pi}{2}\frac{t}{t_p}\right) \\ &= \cos(\theta)I - \mathrm{i}(-1)^{s_i}P_{\boldsymbol{b}_i}\sin(\theta), \end{aligned} \tag{E3}$$

with the Pauli generator $P_{\boldsymbol{b}}$ from Eq. (32) and $\theta(t) := \pi/2(t/t_p)$. Then, the effect of the conjugation on the ith qubit is given by

$$\begin{aligned} S_{c_i}^{-1}(\theta)P_{\boldsymbol{a}_i}S_{c_i}(\theta) &= \underbrace{\left(\cos^2(\theta) + (-1)^{\langle \boldsymbol{a}_i,\boldsymbol{b}_i\rangle}\sin^2(\theta)\right)}_{=:\alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta)}P_{\boldsymbol{a}_i} + (-1)^{s_i}\underbrace{\cos(\theta)\sin(\theta)}_{\beta(\theta)}\mathrm{i}[P_{\boldsymbol{b}_i},P_{\boldsymbol{a}_i}] \\ &= \alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta)P_{\boldsymbol{a}_i} + (-1)^{s_i}\beta(\theta)\mathrm{i}[P_{\boldsymbol{b}_i},P_{\boldsymbol{a}_i}], \end{aligned} \tag{E4}$$

with the binary symplectic form $\langle\cdot,\cdot\rangle$ from Eq. (26) and the commutator $[\cdot,\cdot]$. Let

$$F(e) := \begin{cases} P_{\boldsymbol{a}}\alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta), & e = 0, \\ \mathrm{i}[P_{\boldsymbol{b}_i},P_{\boldsymbol{a}_i}]\beta(\theta), & e = 1. \end{cases} \tag{E5}$$

Inserting H_S in Eq. (E1) and applying $A\otimes(B+C) = A\otimes B + A\otimes C$ yields

$$\begin{aligned} H_{\mathrm{err},\boldsymbol{c}} &= \frac{4t_p}{\pi}\sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} J_{\boldsymbol{a}}\int_0^{\frac{\pi}{2}}\left(\bigotimes_{i\in\mathrm{supp}(\boldsymbol{a})}\left(P_{\boldsymbol{a}_i}\alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta) + \mathrm{i}[P_{\boldsymbol{b}_i},P_{\boldsymbol{a}_i}]\beta(\theta)(-1)^{s_i}\right)\right)\mathrm{d}\theta \\ &= \frac{4t_p}{\pi}\sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}}\left(J_{\boldsymbol{a}}\int_0^{\frac{\pi}{2}}\prod_{i\in\mathrm{supp}(\boldsymbol{a})}\alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta)\mathrm{d}\theta P_{\boldsymbol{a}} + \sum_{\substack{\boldsymbol{e}\in\mathbb{F}_2^n\setminus\{\boldsymbol{0}\} \\ e_i=0\ \forall i\notin\mathrm{supp}(\boldsymbol{a})}}\left(\prod_{i\in\mathrm{supp}(\boldsymbol{a})}(-1)^{e_i+s_i}\right)\left(\bigotimes_{i\in\mathrm{supp}(\boldsymbol{a})}F(e_i)\right)\right) \\ &= \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} J_{\boldsymbol{a}}E_{\boldsymbol{a},\boldsymbol{c}}^{(r\times s)}P_{\boldsymbol{a}} + \sum_{\substack{\boldsymbol{e}\in\mathbb{F}_2^n\setminus\{\boldsymbol{0}\} \\ e_i=0\ \forall i\notin\mathrm{supp}(\boldsymbol{a})}}(-1)^{\boldsymbol{e}\cdot\boldsymbol{s}}\bigotimes_{i\in\mathrm{supp}(\boldsymbol{a})}F(e_i). \end{aligned} \tag{E6}$$

Moreover, we define $E^{(r\times s)}_{\boldsymbol{a},\boldsymbol{c}} := 4t_p/\pi \int_0^{\pi/2} \left(\prod_{i\in\mathrm{supp}(\boldsymbol{a})} \alpha_{\boldsymbol{a}_i,\boldsymbol{b}_i}(\theta)\right) \mathrm{d}\theta$. To conclude, we have the average Hamiltonian

$$H_{\mathrm{av},\boldsymbol{c}}(t) = t\lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}} + H_{\mathrm{err},\boldsymbol{c}}, \tag{E7}$$

from Lemma 4, and we have calculated the finite-pulse-time error term

$$H_{\mathrm{err},\boldsymbol{c}} = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} \left(J_{\boldsymbol{a}} E^{(r\times s)}_{\boldsymbol{a},\boldsymbol{c}} P_{\boldsymbol{a}} + R_{\boldsymbol{a},\boldsymbol{c}}\right), \tag{E8}$$

with

$$E^{(r\times s)}_{\boldsymbol{a},\boldsymbol{c}} := \frac{4t_p}{\pi} \int_0^{\frac{\pi}{2}} \left(\prod_{i\in\mathrm{supp}(\boldsymbol{a})} (\cos^2(\theta) + (-1)^{\langle \boldsymbol{a}_i,\boldsymbol{b}_i\rangle} \sin^2(\theta)) \right) \mathrm{d}\theta \tag{E9}$$

and

$$R_{\boldsymbol{a},\boldsymbol{c}} := \sum_{\substack{\boldsymbol{e}\in\mathbb{F}_2^n\setminus\{\boldsymbol{0}\} \\ e_i=0\ \forall i\notin\mathrm{supp}(\boldsymbol{a})}} (-1)^{\boldsymbol{e}\cdot\boldsymbol{s}} \bigotimes_{i\in\mathrm{supp}(\boldsymbol{a})} F(e_i), \tag{E10}$$

which proofs the lemma. ∎

Proposition 7. The target time evolution $\mathrm{e}^{-\mathrm{i}tH_T}$ can be approximated by a deterministic product formula implementing $U(t\lambda_{\boldsymbol{c}})$, with $\lambda_{\boldsymbol{c}}$ from Eq. (robustPauliLP), and choosing the robust rotation directions $\boldsymbol{s}\in\mathcal{S}_{\boldsymbol{J}}$ of the π pulses. The only approximation errors are given by the ones from the Magnus approximation (Lemma 4) and the approximation error from the deterministic product formula.

Proof. Similarly to the proof of Theorem 5 the columns of $E^{(r\times s)}$ in Eq. (robustPauliLP) have to be rescaled by the number of implementations $U(t\lambda_{\boldsymbol{c}})$ which we denote by $n_{\boldsymbol{c}}$. For simplicity we assume the first-order Trotter approximation in Eq. (18).

We start with the rest term in the average Hamiltonian in Eq. (73) which is proportional to $(-1)^{\boldsymbol{e}_{\boldsymbol{a}}\cdot\boldsymbol{s}}$. Let us assume that we require κ different rotation directions $\boldsymbol{s}_{(1)},\ldots,\boldsymbol{s}_{(\kappa)}\in\mathbb{F}_2^n$ such that $\sum_{j=1}^{\kappa}(-1)^{\boldsymbol{e}_{\boldsymbol{a}}\cdot\boldsymbol{s}_{(j)}} = 0$ for all $\boldsymbol{e}_{\boldsymbol{a}}\in\mathbb{F}_2^n$ with $e_{\boldsymbol{a},i}=0$ for all $i\notin\mathrm{supp}(\boldsymbol{a})$ for any $\boldsymbol{a}$ with $J_{\boldsymbol{a}}\neq 0$. A nonoptimal choice of $\boldsymbol{s}_{(j)}$ is all possible binary combinations without the all-zero vector; then $\kappa = 2^n - 1$. An efficient choice is provided in Proposition 8 below. To indicate the used rotation direction, we specify the single-qubit Pauli pulse layer by the tuple $\boldsymbol{c}_{(j)} = (t_p, \boldsymbol{s}_{(j)}, \boldsymbol{h})$. Note that the generators do not change. Let the partial average Hamiltonian be the average Hamiltonian without the rest error term,

$$\begin{aligned}\tilde{H}_{\mathrm{av},\boldsymbol{c}}(t) &:= \sum_{j=1}^{\kappa} H_{\mathrm{av},\boldsymbol{c}_{(j)}}\left(\frac{t}{\kappa}\right) \\ &= t\lambda_{\boldsymbol{c}} \boldsymbol{S}_{\boldsymbol{c}}^\dagger H_S \boldsymbol{S}_{\boldsymbol{c}} + \kappa \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} E^{(r\times s)}_{\boldsymbol{a},\boldsymbol{c}} J_{\boldsymbol{a}} P_{\boldsymbol{a}}.\end{aligned} \tag{E11}$$

The evolution under the partial average Hamiltonian can be approximated by

$$\mathrm{e}^{-\mathrm{i}t\tilde{H}_{\mathrm{av},\boldsymbol{c}}(t)} \approx \prod_{j=1}^{\kappa} \mathrm{e}^{-\mathrm{i}H_{\mathrm{av},\boldsymbol{c}_{(j)}}\left(\frac{t}{\kappa}\right)} \approx \prod_{j=1}^{\kappa} U(t\lambda_{\boldsymbol{c}_{(j)}}/\kappa), \tag{E12}$$

where the first approximation is given by a first-order Trotter scheme where we set $n_{\mathrm{Tro}} = 1$ for simplicity, and the second approximation is given by the first-order Magnus expansion from Lemma 4. Let $t\tilde{H}_{\mathrm{av}} := \sum_{\boldsymbol{c}}^{s} \tilde{H}_{\mathrm{av},\boldsymbol{c}}(t)$, with $\lambda_{\boldsymbol{c}}$ from Eq. (robustPauliLP). Let the target Hamiltonian be

$$H_T = \sum_{\boldsymbol{a}\in\mathbb{F}_2^{2n}\setminus\{\boldsymbol{0}\}} A_{\boldsymbol{a}} P_{\boldsymbol{a}}. \tag{E13}$$

By the constraint of Eq. (robustPauliLP), we have $H_{\mathrm{av}} = H_T$. The time evolution governed by the target Hamiltonian H_T can be approximated with

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}tH_T} = \mathrm{e}^{-\mathrm{i}tH_{\mathrm{av}}} &\approx \left(\prod_{\boldsymbol{c}} \mathrm{e}^{-\mathrm{i}\tilde{H}_{\mathrm{av},\boldsymbol{c}_{(j)}}\left(\frac{t}{n_{\mathrm{Tro}}}\right)}\right)^{n_{\mathrm{Tro}}} \\ &\approx \left(\prod_{\boldsymbol{c}} \prod_{j=1}^{\kappa} U\left(\frac{t\lambda_{\boldsymbol{c}_{(j)}}}{\kappa n_{\mathrm{Tro}}}\right)\right)^{n_{\mathrm{Tro}}},\end{aligned} \tag{E14}$$

where the first approximation is given by the Trotter scheme. To account for the number of implemented Pauli conjugations $U(t\lambda_{\boldsymbol{c}})$ we have to rescale $E^{(r\times s)}$ by $n_{\boldsymbol{c}} = \kappa n_{\mathrm{Tro}}$ for all $\boldsymbol{c}$. ∎

Proposition E1. Let the rotation directions $\boldsymbol{s}\in\mathbb{F}_2^n$ of the π pulses such that $\sum_{\boldsymbol{s}}(-1)^{\boldsymbol{e}_{\boldsymbol{a}}\cdot\boldsymbol{s}} = 0$ for all $\boldsymbol{e}_{\boldsymbol{a}}\in\mathbb{F}_2^n$ with $e_{\boldsymbol{a},i}=0$ for all $i\notin\mathrm{supp}(\boldsymbol{a})$ for any $\boldsymbol{a}$ with $J_{\boldsymbol{a}}\neq 0$, as in Lemma 6. Then, the rotation-angle errors in the first-order Taylor approximation cancel.

Proof. We model the rotation-angle error of a perfect π pulse $\mathrm{e}^{-\mathrm{i}(-1)^{s_i}(\pi/2)P_{\boldsymbol{b}_i}}$ with the rotation direction $(-1)^{s_i}$ and $s_i\in\mathbb{F}_2$ by

$$\tilde{S}_{c_i} = \mathrm{e}^{-\mathrm{i}(-1)^{s_i}\frac{\pi+\varepsilon_i}{2}P_{\boldsymbol{b}_i}} = -\frac{\varepsilon_i}{2}\,\mathrm{Id} - \mathrm{i}(-1)^{s_i}P_{\boldsymbol{b}_i} + \mathrm{O}(\varepsilon_i^2), \tag{E15}$$

where we have used the first-order Taylor expansion of sine and cosine. Conjugating a Pauli operator with an

imperfect Pauli pulse results in

$$\tilde{S}_{c_i}^\dagger P_{\boldsymbol{a}_i} \tilde{S}_{c_i} = P_{\boldsymbol{b}_i} P_{\boldsymbol{a}_i} P_{\boldsymbol{b}_i} + \mathrm{i}(-1)^{s_i} \frac{\varepsilon_i}{2} [P_{\boldsymbol{a}_i}, P_{\boldsymbol{b}_i}] + \mathrm{O}(\varepsilon_i^2). \tag{E16}$$

We use the identity $U^\dagger \mathrm{e}^{-\mathrm{i}tH} U = \mathrm{e}^{-\mathrm{i}tU^\dagger HU}$ to compute the effective Hamiltonian of a system Hamiltonian H_S conjugated with a layer of imperfect Pauli pulses $\tilde{\boldsymbol{S}}_{\boldsymbol{c}} = \bigotimes_{i=1}^n \tilde{S}_{c_i}$,

$$\begin{aligned}
\tilde{\boldsymbol{S}}_{\boldsymbol{c}}^\dagger H_S \tilde{\boldsymbol{S}}_{\boldsymbol{c}} &= \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} J_{\boldsymbol{a}} \bigotimes_{i=1}^n \left(\tilde{S}_{c_i}^\dagger P_{\boldsymbol{a}_i} \tilde{S}_{c_i} \right) \\
&= \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} J_{\boldsymbol{a}} \bigotimes_{i=1}^n \left(P_{\boldsymbol{b}_i} P_{\boldsymbol{a}_i} P_{\boldsymbol{b}_i} + \mathrm{i}(-1)^{s_i} \frac{\varepsilon_i}{2} [P_{\boldsymbol{a}_i}, P_{\boldsymbol{b}_i}] + \mathrm{O}(\varepsilon_i^2) \right) \\
&= \sum_{\boldsymbol{a} \in \mathbb{F}_2^{2n} \setminus \{\boldsymbol{0}\}} \left(J_{\boldsymbol{a}} P_{\boldsymbol{b}} P_{\boldsymbol{a}} P_{\boldsymbol{b}} + \sum_{\substack{\boldsymbol{e} \in \mathbb{F}_2^n \\ e_i = 0 \ \forall i \notin \mathrm{supp}(\boldsymbol{a})}} \left(\prod_{i \in \mathrm{supp}(\boldsymbol{a})} (-1)^{e_i + s_i} \varepsilon_i^{e_i} \right) \left(\bigotimes_{i \in \mathrm{supp}(\boldsymbol{a})} F(e_i) + \mathrm{O}(\varepsilon_i^2) \right) \right),
\end{aligned} \tag{E17}$$

where we have used $A \otimes (B + C) = A \otimes B + A \otimes C$ and

$$F(e) := \begin{cases} P_{\boldsymbol{b}_i} P_{\boldsymbol{a}_i} P_{\boldsymbol{b}_i}, & e = 0 \\ \mathrm{i} \frac{1}{2} [P_{\boldsymbol{a}_i}, P_{\boldsymbol{b}_i}], & e = 1. \end{cases} \tag{E18}$$

The second term in the sum is the first-order angle error contribution, and has the same sign structure $(-1)^{e_i + s_i}$ as the rest term $R_{\boldsymbol{a},\boldsymbol{c}}$ in Eq. (E6). Therefore, choosing the signs $\boldsymbol{s} \in \mathbb{F}_2^n$, such that the rest term $R_{\boldsymbol{a},\boldsymbol{c}}$ cancels, simultaneously cancels the first-order angle error contribution in Eq. (E17). ■

The proof of the following result is similar to the proof in Ref. [42, Lemma 8].

Proposition 8. Let $\kappa = 2^{\lceil \log_2(n+1) \rceil} \leq 2n$ and let $W^{(\kappa \times \kappa)}$ be the $\kappa \times \kappa$ dimensional Walsh-Hadamard matrix. Choose n distinct columns from $W^{(\kappa \times \kappa)}$ without the first column and define the resulting partial Walsh-Hadamard matrix as $W^{(\kappa \times n)}$. Let $(-1)^{\boldsymbol{s}_{(j)}}$ be the jth row of $W^{(\kappa \times n)}$. Then, for any nonzero two-body interaction $J_{\boldsymbol{a}} \neq 0$ with $|\mathrm{supp}(\boldsymbol{a})| = 2$ we have $\boldsymbol{s}_{(j)} \in \mathcal{S}_{\boldsymbol{J}}$ for all $j = 1, \ldots, \kappa$.

Proof. From the choice of $(-1)^{\boldsymbol{s}_{(j)}}$ it directly follows that $\sum_{j=1}^\kappa (-1)^{\boldsymbol{e} \cdot \boldsymbol{s}_{(j)}} = 0$ for all $\boldsymbol{e} \in \mathbb{F}_2^n$ with $|\boldsymbol{e}| = 1$, since the sum over all rows of a Walsh-Hadamard matrix, especially $W^{(\kappa \times n)}$, is zero. We now have to show that $\sum_{j=1}^\kappa (-1)^{\boldsymbol{e} \cdot \boldsymbol{s}_{(j)}} = 0$ for all $\boldsymbol{e} \in \mathbb{F}_2^n$ with $|\boldsymbol{e}| = 2$. For any $\boldsymbol{e} \in \mathbb{F}_2^n$ with $|\boldsymbol{e}| = 2$ we can write $(-1)^{\boldsymbol{e} \cdot \boldsymbol{s}} = (-1)^{s_i} (-1)^{s_k} = \left(((-1)^{\boldsymbol{s}}) ((-1)^{\boldsymbol{s}})^T \right)_{ik}$ with $i, k \in [n]$ and $i \neq k$. Then, we obtain

$$\sum_{j=1}^\kappa (-1)^{\boldsymbol{e} \cdot \boldsymbol{s}_{(j)}} = \sum_{j=1}^\kappa \left(((-1)^{\boldsymbol{s}_{(j)}}) ((-1)^{\boldsymbol{s}_{(j)}})^T \right)_{ik}, \tag{E19}$$

with $i, k \in [n]$ and $i \neq k$. The orthogonality property of the Walsh-Hadamard matrix yields

$$\sum_{j=1}^\kappa ((-1)^{\boldsymbol{s}_{(j)}}) ((-1)^{\boldsymbol{s}_{(j)}})^T = (W^{(\kappa \times n)})^T W^{(\kappa \times n)} = \kappa I, \tag{E20}$$

where the nondiagonal entries $i \neq k$ correspond to the sum in Eq. (E19) and are zero. ■

[1] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Simulating physical phenomena by quantum networks, Phys. Rev. A **65**, 042323 (2002).

[2] R. Blatt and C. F. Roos, Quantum simulations with trapped ions, Nat. Phys. **8**, 277 (2012).

[3] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Solving strongly correlated electron models on a quantum computer, Phys. Rev. A **92**, 062318 (2015).

[4] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Probing many-body dynamics on a 51-atom quantum simulator, Nature **551**, 579 (2017).

[5] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature **551**, 601 (2017).
[6] N. Defenu, A. Lerose, and S. Pappalardi, Out-of-equilibrium dynamics of quantum many-body systems with long-range interactions, Phys. Rep. **1074**, 1 (2024).
[7] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Simulating chemistry using quantum computers, Annu. Rev. Phys. Chem. **62**, 185 (2011).
[8] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Gate-count estimates for performing quantum chemistry on small quantum computers, Phys. Rev. A **90**, 022305 (2014).
[9] J. Olson, Y. Cao, J. Romero, P. Johnson, P.-L. Dallaire-Demers, N. Sawaya, P. Narang, I. Kivlichan, M. Wasielewski, and A. Aspuru-Guzik, Quantum information and computation for chemistry, ArXiv:1706.05413.
[10] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, Proc. Natl. Acad. Sci. **115**, 9456 (2018).
[11] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, Nature **607**, 667 (2022).
[12] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).
[13] R. Trivedi, A. Franco Rubio, and J. I. Cirac, Quantum advantage and stability to errors in analogue quantum simulators, Nat. Commun. **15**, 6507 (2024).
[14] D. W. Leung, I. L. Chuang, F. Yamaguchi, and Y. Yamamoto, Efficient implementation of coupled logic gates for quantum computation, Phys. Rev. A **61**, 042310 (2000).
[15] D. Leung, Simulation and reversal of n-qubit Hamiltonians using Hadamard matrices, J. Mod. Opt. **49**, 1199 (2002).
[16] J. L. Dodd, M. A. Nielsen, M. J. Bremner, and R. T. Thew, Universal quantum computation and simulation using any entangling Hamiltonian and local unitaries, Phys. Rev. A **65**, 040301(R) (2002).
[17] M. A. Nielsen, M. J. Bremner, J. L. Dodd, A. M. Childs, and C. M. Dawson, Universal simulation of Hamiltonian dynamics for quantum systems with finite-dimensional state spaces, Phys. Rev. A **66**, 022317 (2002).
[18] M. Votto, J. Zeiher, and B. Vermersch, Universal quantum processors in spin systems via robust local pulse sequences, Quantum **8**, 1513 (2024).
[19] J. Choi, H. Zhou, H. S. Knowles, R. Landig, S. Choi, and M. D. Lukin, Robust dynamic Hamiltonian engineering of many-body spin systems, Phys. Rev. X **10**, 031002 (2020).
[20] S. Choi, N. Y. Yao, and M. D. Lukin, Dynamical engineering of interactions in qudit ensembles, Phys. Rev. Lett. **119**, 183603 (2017).
[21] H. Zhou, H. Gao, N. T. Leitao, O. Makarova, I. Cong, A. M. Douglas, L. S. Martin, and M. D. Lukin, Robust Hamiltonian engineering for interacting qudit systems, Phys. Rev. X **14**, 031017 (2024).
[22] S. Geier, N. Thaicharoen, C. Hainaut, T. Franz, A. Salzinger, A. Tebben, D. Grimshandl, G. Zürn, and M. Weidemüller, Floquet Hamiltonian engineering of an isolated many-body spin system, Science **374**, 1149 (2021).
[23] H. Zhou, J. Choi, S. Choi, R. Landig, A. M. Douglas, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, P. Cappellaro, H. S. Knowles, H. Park, and M. D. Lukin, Quantum metrology with strongly interacting spin systems, Phys. Rev. X **10**, 031003 (2020).
[24] P. Scholl, H. J. Williams, G. Bornet, F. Wallner, D. Barredo, L. Henriet, A. Signoles, C. Hainaut, T. Franz, S. Geier, A. Tebben, A. Salzinger, G. Zürn, T. Lahaye, M. Weidemüller, and A. Browaeys, Microwave engineering of programmable XXZ Hamiltonians in arrays of Rydberg atoms, PRX Quantum **3**, 020303 (2022).
[25] D. Barredo, H. Labuhn, S. Ravets, T. Lahaye, A. Browaeys, and C. S. Adams, Coherent excitation transfer in a spin chain of three Rydberg atoms, Phys. Rev. Lett. **114**, 113002 (2015).
[26] D. Bluvstein, H. Levine, G. Semeghini, T. T. Wang, S. Ebadi, M. Kalinowski, A. Keesling, N. Maskara, H. Pichler, M. Greiner, V. Vuletić, and M. D. Lukin, A quantum processor based on coherent transport of entangled atom arrays, Nature **604**, 451 (2022).
[27] D. Bluvstein, *et al.*, Logical quantum processor based on reconfigurable atom arrays, Nature **626**, 58 (2023).
[28] M. Garcia-de-Andoin, A. Saiz, P. Pérez-Fernández, L. Lamata, I. Oregi, and M. Sanz, Digital-analog quantum computation with arbitrary two-body Hamiltonians, Phys. Rev. Res. **6**, 013280 (2024).
[29] D. Hayes, S. T. Flammia, and M. J. Biercuk, Programmable quantum simulation by dynamic Hamiltonian engineering, New J. Phys. **16**, 083027 (2014).
[30] 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).
[31] Y. Lu, S. Zhang, K. Zhang, W. Chen, Y. Shen, J. Zhang, J.-N. Zhang, and K. Kim, Global entangling gates on arbitrary ion qubits, Nature **572**, 363 (2019).
[32] N. Grzesiak, R. Blümel, K. Wright, K. M. Beck, N. C. Pisenti, M. Li, V. Chaplin, J. M. Amini, S. Debnath, J.-S. Chen, and Y. Nam, Efficient arbitrary simultaneously entangling gates on a trapped-ion quantum computer, Nat. Commun. **11**, 2963 (2020).
[33] J. K. Pachos and M. B. Plenio, Three-spin interactions in optical lattices and criticality in cluster Hamiltonians, Phys. Rev. Lett. **93**, 056402 (2004).
[34] H. P. Büchler, A. Micheli, and P. Zoller, Three-body interactions with cold polar molecules, Nat. Phys. **3**, 726 (2007).
[35] X. Peng, J. Zhang, J. Du, and D. Suter, Quantum simulation of a system with competing two- and three-body interactions, Phys. Rev. Lett. **103**, 140501 (2009).
[36] K. Zhang, H. Li, P. Zhang, J. Yuan, J. Chen, W. Ren, Z. Wang, C. Song, D.-W. Wang, H. Wang, S. Zhu, G. S. Agarwal, and M. O. Scully, Synthesizing five-body interaction in a superconducting quantum circuit, Phys. Rev. Lett. **128**, 190502 (2022).
[37] H.-Y. Huang, Y. Tong, D. Fang, and Y. Su, Learning many-body Hamiltonians with Heisenberg-limited scaling, Phys. Rev. Lett. **130**, 200403 (2023).
[38] M. Ma, S. T. Flammia, J. Preskill, and Y. Tong, Learning k-body Hamiltonians via compressed sensing, ArXiv:2410.18928.

[39] H.-Y. Hu, M. Ma, W. Gong, Q. Ye, Y. Tong, S. T. Flammia, and S. F. Yelin, Ansatz-free Hamiltonian learning with Heisenberg-limited scaling, PRX Quantum **6**, 040315 (2025).
[40] E. Bach and S. Huiberts, Optimal smoothed analysis of the simplex method, ArXiv:2504.04197.
[41] 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).
[42] P. Baßler, M. Heinrich, and M. Kliesch, Time-optimal multi-qubit gates: Complexity, efficient heuristic and gate-time bounds, Quantum **8**, 1279 (2024).
[43] G. Bhole, T. Tsunoda, P. J. Leek, and J. A. Jones, Rescaling interactions for quantum control, Phys. Rev. Appl. **13**, 034002 (2020).
[44] T. Tsunoda, G. Bhole, S. A. Jones, J. A. Jones, and P. J. Leek, Efficient Hamiltonian programming in qubit arrays with nearest-neighbor couplings, Phys. Rev. A **102**, 032405 (2020).
[45] A. D. Leu, M. F. Gely, M. A. Weber, M. C. Smith, D. P. Nadlinger, and D. M. Lucas, Fast, high-fidelity addressed single-qubit gates using efficient composite pulse sequences, Phys. Rev. Lett. **131**, 120601 (2023).
[46] S. P. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, United Kingdom, 2004).
[47] I. Bárány, A generalization of Carathéodory's theorem, Discrete Math. **40**, 141 (1982).
[48] D. A. Spielman and S.-H. Teng, Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time, J. ACM **51**, 385 (2004).
[49] M. Bando, T. Ichikawa, Y. Kondo, and M. Nakahara, Concatenated composite pulses compensating simultaneous systematic errors, J. Phys. Soc. Jpn. **82**, 014004 (2013).
[50] S. Kukita, H. Kiya, and Y. Kondo, Short composite quantum gate robust against two common systematic errors, J. Phys. Soc. Jpn. **91**, 104001 (2022).
[51] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, A rewriting system for convex optimization problems, J. Control Decis. **5**, 42 (2018).
[52] S. Diamond and S. Boyd, CVXPY: A PYTHON-embedded modeling language for convex optimization, J. Mach. Learn. Res. **17**, 1 (2016).
[53] MOSEK ApS, MOSEK *Optimizer API for* PYTHON *9.3.14* (2022), https://docs.mosek.com/latest/pythonapi/index.html.
[54] P. Baßler, Source code for "Efficient Hamiltonian engineering," https://github.com/paba92/EffHamEng, 2024.
[55] F. Arute, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).
[56] L. Schmid, D. F. Locher, M. Rispler, S. Blatt, J. Zeiher, M. Müller, and R. Wille, Computational capabilities and compiler development for neutral atom quantum processors—connecting tool developers and hardware experts, Quantum Sci. Technol. **9**, 033001 (2024).
[57] L. Henriet, L. Beguin, A. Signoles, T. Lahaye, A. Browaeys, G.-O. Reymond, and C. Jurczak, Quantum computing with neutral atoms, Quantum **4**, 327 (2020).
[58] H. K. Cummins, G. Llewellyn, and J. A. Jones, Tackling systematic errors in quantum logic gates with composite rotations, Phys. Rev. A **67**, 042308 (2003).
[59] C. Piltz, T. Sriarunothai, S. S. Ivanov, S. Wölk, and C. Wunderlich, Versatile microwave-driven trapped ion spin system for quantum information processing, Sci. Adv. **2**, e1600093 (2016).
[60] J. Farkas, Theorie der einfachen Ungleichungen, J. Reine Angew. Math. **124**, 1 (1902).
[61] E. Stiemke, Über positive Lösungen homogener linearer Gleichungen, Math. Ann. **76**, 340 (1915).
[62] S. Hayakawa, T. Lyons, and H. Oberhauser, Estimating the probability that a given vector is in the convex hull of a random sample, Probab. Theory Relat. Fields **185**, 705 (2023).
[63] J. G. Wendel, A problem in geometric probability, Math. Scand. **11**, 109 (1962).
[64] U. Wagner and E. Welzl, A continuous analogue of the upper bound theorem, Discrete Comput. Geom. **26**, 205 (2001).
[65] V. P. Canelles, M. G. Algaba, H. Heimonen, M. Papič, M. Ponce, J. Rönkkö, M. J. Thapa, I. de Vega, and A. Auer, Benchmarking digital-analog quantum computation, ArXiv:2307.07335.
[66] U. Haeberlen and J. S. Waugh, Coherent averaging effects in magnetic resonance, Phys. Rev. **175**, 453 (1968).
[67] U. Haeberlen, *High Resolution NMR in Solids* (Academic Press, New York, San Francisco, London, 1976).
[68] W. Magnus, On the exponential solution of differential equations for a linear operator, Commun. Pure Appl. Math. **7**, 649 (1954).
[69] B. T. Torosov and N. V. Vitanov, Composite pulses with errant phases, Phys. Rev. A **100**, 023410 (2019).
[70] G. T. Genov, D. Schraft, T. Halfmann, and N. V. Vitanov, Correction of arbitrary field errors in population inversion of quantum systems by universal composite pulses, Phys. Rev. Lett. **113**, 043001 (2014).
[71] G. T. Genov, M. Hain, N. V. Vitanov, and T. Halfmann, Universal composite pulses for efficient population inversion with an arbitrary excitation profile, Phys. Rev. A **101**, 013827 (2020).
[72] H.-N. Wu, C. Zhang, J. Song, Y. Xia, and Z.-C. Shi, Composite pulses for optimal robust control in two-level systems, Phys. Rev. A **107**, 023103 (2023).
[73] C. Kabytayev, T. J. Green, K. Khodjasteh, M. J. Biercuk, L. Viola, and K. R. Brown, Robustness of composite pulses to time-dependent control noise, Phys. Rev. A **90**, 012316 (2014).
[74] B. T. Torosov, S. S. Ivanov, and N. V. Vitanov, Narrowband and passband composite pulses for variable rotations, Phys. Rev. A **102**, 013105 (2020).
[75] P. García-Molina, A. Martin, M. Garcia de Andoin, and M. Salz, Mitigating noise in digital and digital–analog quantum computation, Commun. Phys. **7**, 321 (2024).
[76] M. E. S. Morales, P. C. S. Costa, G. Pantaleoni, D. K. Burgarth, Y. R. Sanders, and D. W. Berry, Selection and improvement of product formulae for best performance of quantum simulation, Quantum Inf. Comput. **25**, 1 (2025).
[77] H. F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. **10**, 545 (1959).
[78] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, J. Math. Phys. **32**, 400 (1991).
[79] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, Phys. Rev. X **11**, 011020 (2021).

[80] P. C. Moan and J. Niesen, Convergence of the Magnus series, Found. Comput. Math. **8**, 291 (2008).

[81] A. Brinkmann, Introduction to average Hamiltonian theory. I. Basics, Concepts Magn. Reson. Part A **45A**, e21414 (2016).

[82] K. Sharma and M. C. Tran, Hamiltonian simulation in the interaction picture using the Magnus expansion, ArXiv:2404.02966.

[83] N. B. Karahanoğlu, H. Erdoğan, and Ş. İ. Birbil, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, Piscataway, NJ, USA, 2013), p. 5870.

[84] A. Arnal, F. Casas, and C. Chiralt, A general formula for the Magnus expansion in terms of iterated integrals of right-nested commutators, J. Phys. Commun. **2**, 035024 (2018).