Supplemental Material: Fixed Depth Hamiltonian Simulation via Cartan Decomposition

Efekan Kökcü,[1,*] Thomas Steckmann,[1] Yan Wang,[2] J. K. Freericks,[3]
Eugene F. Dumitrescu,[2,†] and Alexander F. Kemper[1,‡]

[1] *Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA*
[2] *Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[3] *Department of Physics, Georgetown University, 37th and O Sts. NW, Washington, DC 20057 USA*

(Dated: June 13, 2022)

CONTENTS

I. HAMILTONIAN ALGEBRAS OF CERTAIN MODELS

A. XY Model

For the 1-D nearest neighbour XY model with open boundary conditions and arbitrary interaction coefficients,

$$\mathcal{H} = \sum_{i=1}^{n-1} (a_i X_i X_{i+1} + b_i Y_i Y_{i+1}), \tag{S.1}$$

and the Hamiltonian algebra is found to be

$$\mathfrak{g}(\mathrm{XY}) = \mathrm{span}\{\widehat{X_i X}_{i+a}, \widehat{Y_i Y}_{i+a}, \widehat{X_i Y}_{i+b}, \widehat{Y_i X}_{i+b} \big| a \text{ odd}, b \text{ even}, 1 \leq i, i+a, i+b \leq n\}. \tag{S.2}$$

The dimension of this algebra is calculated as $|\mathfrak{g}(\mathrm{XY})| = 2\binom{n}{2} = n(n-1)$.

B. TFIM and TFXY Model

For the 1-D nearest neighbour transverse field XY model with open boundary conditions and free coefficients,

$$\mathcal{H} = \sum_{i=1}^{n-1} (a_i X_i X_{i+1} + b_i Y_i Y_{i+1}) + \sum_{i=1}^{n} c_i Z_i, \tag{S.3}$$

* ekokcu@ncsu.edu
† dumitrescuef@ornl.gov
‡ akemper@ncsu.edu

the Hamiltonian algebra is found to be

$$\mathfrak{g}(\text{TFXY}) = \text{span}\{Z_j, \widehat{X_iX}_j, \widehat{Y_iY}_j, \widehat{X_iY}_j, \widehat{Y_iX}_j \big| 1 \leq i,j \leq n;\ i<j\}. \tag{S.4}$$

The same algebra is found for transverse field Ising model, i.e. the $b_i = 0$ case for the Hamiltonian given in (S.3). The dimension of this algebra is $|\mathfrak{g}(\text{TFXY})| = n + 4\binom{n}{2} = n(2n-1)$.

C. Heisenberg Model

For the 1-D nearest neighbour Heisenberg model with open boundary conditions and free coefficients,

$$\mathcal{H} = \sum_{i=1}^{n-1}(a_iX_iX_{i+1} + b_iY_iY_{i+1} + c_iZ_iZ_{i+1}), \tag{S.5}$$

the Hamiltonian algebra is found to be

$$\mathfrak{g}(\text{Heisenberg}) = \text{span}\Big(\{\text{Pauli strings with } a \text{ many } X,\ b \text{ many } Y,\ c \text{ many } Z \big| a+b, a+c, b+c \text{ even},\ a,b,c \geq 1\} \\ \setminus \{XXX...X,\ YYY...Y,\ ZZZ...Z\}\Big) \tag{S.6}$$

All basis elements in this algebra commute with $XXX...X$, $YYY...Y$, and therefore also commute with $ZZZ...Z$. Any other Pauli string apart from the ones in algebra does not commute either with $XXX...X$ or $YYY...Y$.

In order to determine the dimension of $\mathfrak{g}(\text{Heisenberg})$, let us decompose $\mathfrak{su}(2^n) = \mathfrak{k} \oplus \mathfrak{m}$ with $\theta(g) = XXX...X\, g\, XXX...X$. Then $\mathfrak{k}$ is the subalgebra of $\mathfrak{su}(2^n)$ consisting of all the elements in $\mathfrak{su}(2^n)$ that commute with $XXX...X$. This decomposition of $\mathfrak{su}(2^n)$ is type A III, and resulting $\mathfrak{k}$ is isomorphic to $\mathfrak{k} \cong \mathfrak{su}(2^{n-1}) \oplus \mathfrak{su}(2^{n-1}) \oplus \mathfrak{u}(1)$ [1, 2]. To have all the elements that commute both with $XXX...X$ and $YYY...Y$, let us further decompose $\mathfrak{k}$ into $\mathfrak{k} = \mathfrak{k}' \oplus \mathfrak{m}'$ with the involution $\theta'(g) = YYY...Y\, g\, YYY...Y$. Therefore we have

$$\mathfrak{k}' = \text{span}\{\text{Pauli strings with } a \text{ many } X,\ b \text{ many } Y,\ c \text{ many } Z \big| a+b, a+c, b+c \text{ even},\ a,b,c \geq 1\}. \tag{S.7}$$

This decomposition does not affect $\mathfrak{u}(1)$ component of $\mathfrak{k}$. It only affects $\mathfrak{su}(2^{n-1})$ pieces separately, and leads to $\mathfrak{k}' \cong (\mathfrak{su}(2^{n-2}) \oplus \mathfrak{su}(2^{n-2}) \oplus \mathfrak{u}(1)) \oplus (\mathfrak{su}(2^{n-2}) \oplus \mathfrak{su}(2^{n-2}) \oplus \mathfrak{u}(1)) \oplus \mathfrak{u}(1)$. Therefore the dimension of $\mathfrak{k}'$ is $|\mathfrak{k}'| = 4|\mathfrak{su}(2^{n-2})| + 3 = 4^{n-1} - 1$.

The difference between the basis of $\mathfrak{k}'$ and $\mathfrak{g}(\text{Heisenberg})$ are the elements $XXX...X$, $YYY...Y$, and $ZZZ...Z$. Therefore the dimension of the Heisenberg Hamiltonian algebra can be calculated as $|\mathfrak{g}(\text{Heisenberg})| = |\mathfrak{k}'| - 3 = 4^{n-1} - 4$.

II. REVIEW OF INVOLUTION

A. Involution and Cartan Decomposition

The Cartan decomposition is defined as the split given in Def. 1, i.e. $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$. Finding such a split is difficult; assigning some basis elements into $\mathfrak{k}$ and some into $\mathfrak{m}$ may lead to inconsistencies as we go through the basis elements of the full algebra $\mathfrak{g}$.

To get around this issue we may use an involution θ, i.e. a homomorphism on $\mathfrak{g}$ that preserves commutation relations and satisfies $\theta(\theta(g)) = g$ for all $g \in \mathfrak{g}$. This naturally splits the algebra when one considers its $+1$ and -1 eigen-solutions (since the square of the involution is identity, $+1$ and -1 are the only possibilities for eigenvalues). Let us name p_i as a $+1$ eigen-solutions, and n_i as a -1 eigen-solutions where $i = 1, 2, 3, ...$ such that $\theta(p_i) = p_i$ and $\theta(n_i) = -n_i$. Then we can see that $[p_i, p_j]$ is a $+1$ eigen-solution:

$$\theta([p_i, p_j]) = [\theta(p_i), \theta(p_j)] = [p_i, p_j], \tag{S.8}$$

$[p_i, n_j]$ is a -1 eigen-solution:

$$\theta([p_i, n_j]) = [\theta(p_i), \theta(n_j)] = [p_i, -n_j] = -[p_i, n_j], \tag{S.9}$$

and $[n_i, n_j]$ is a $+1$ eigen-solution:

$$\theta([n_i, n_j]) = [\theta(n_i), \theta(n_j)] = [-n_i, -n_j] = [n_i, n_j]. \tag{S.10}$$

Therefore if one defines $\mathfrak{k}$ as the positive eigen-solution space for the involution θ and $\mathfrak{m}$ as the negative eigen-solution space, then the commutation relations given in Def. 1 are automatically satisfied.

B. Involution Types for $\mathfrak{su}(N)$

The unitary operators involved in quantum computing fall into the Lie group $SU(N)$ which is generated by the Lie algebra $\mathfrak{su}(N)$, where $N = 2^n$ and n is the number of qubits. This algebra has infinitely many Cartan decompositions, many of them are equivalent to each other after a similarity transformation. However, there are 3 different classes of Cartan decompositions of $\mathfrak{su}(N)$ that cannot be transformed to each other which are named as AI, AII and AIII [3, 4]. Involutions that correspond to these (up to a similarity transformation) are as follows.

- For AI, $\theta(g) = -g^T$ (In [3] this involution is given as complex conjugation rather than transpose, which is equivalent due to the fact that all the matrices we work are Hermitian matrices in this paper.)
- For AII (only when N is even), $\theta(g) = -Mg^TM$, where

$$M = \begin{pmatrix} 0 & -I_{N/2} \\ I_{N/2} & 0 \end{pmatrix} \tag{S.11}$$

  and $I_{N/2}$ is the $N/2 \times N/2$ identity matrix
- For AIII, $\theta(g) = UgU$ where U is a diagonal matrix that has each diagonal element to be either $+1$ or -1 (note that $U^T = U$).

C. Involution for Certain Models

Pauli matrices satisfy $X^T = X$, $Y^T = -Y$, $Z^T = Z$. Using $(A \otimes B)^T = A^T \otimes B^T$, one finds that Pauli strings with an even number of Y matrices satisfy $\sigma^T = \sigma$, while the ones with an odd number of Y matrices satisfy $\sigma^T = -\sigma$. Therefore, for any Lie subalgera $\mathfrak{g} \subseteq \mathfrak{su}(2^n)$ using the involution $\theta(g) = -g^T$ leads to a Cartan decomposition via $\theta(\mathfrak{k}) = \mathfrak{k}$, $\theta(\mathfrak{m}) = -\mathfrak{m}$ as

$$\begin{aligned} \mathfrak{k} &= \text{span}\{\text{Pauli strings} \in \mathfrak{g} \text{ with odd } Y \text{ matrices}\}, \\ \mathfrak{m} &= \text{span}\{\text{Pauli strings} \in \mathfrak{g} \text{ with even } Y \text{ matrices}\}. \end{aligned} \tag{S.12}$$

The XY, transverse field XY, transverse field Ising and Heisenberg models have Hamiltonians consisting only of Pauli strings that have either 0 or 2 Y matrices, therefore satisfy $\theta(\mathcal{H}) = -\mathcal{H}^T = -\mathcal{H}$, which makes $\theta(g) = -g^T$ a suitable involution to apply Theorem 1 to these models.

After the Jordan Wigner transformation, all these models fall into a fermion model that has time reversal symmetry. The reason that the involution $\theta(g) = -g^T$ work for these models can be explained by the relation between Cartan decompositions and discrete quantum symmetries as discussed in Ref. [4]. There, it is explained that if the symmetry is due to a unitary transformation, then corresponding involution puts anything symmetric into $\mathfrak{k}$. Parity in space can be considered as one of those. However, if the symmetry is due to an anti-unitary transformation such as time reversal, then the corresponding involution puts the symmetric element into $\mathfrak{m}$. Therefore it is guaranteed that there is an involution that will put $\mathcal{H}$ into $\mathfrak{m}$ if the system has time reversal symmetry.

D. Involution Search

In our algorithm, we are limiting ourselves to single Pauli strings due to circuit composition considerations. This limits our involution pool for all AI, AII and AIII types, and leads to the following list:

1. $\theta(g) = -Bg^TB$ as AI, where B is a Pauli string containing even number of Y matrices,
2. $\theta(g) = -Bg^TB$ as AII, where B is a Pauli string containing odd number of Y matrices,
3. $\theta(g) = BgB$ as AIII, where B is a Pauli string.

One can search for B for any given Hamiltonian $\mathcal{H}$ to find a suitable involution satisfying $\theta(\mathcal{H}) = -\mathcal{H}$.

III. METHOD TO FIND A CARTAN SUBALGEBRA

In this work, we specifically deal with basis elements that are single Pauli strings, i.e. tensor product of n Pauli matrices where n is the number of spins/qubits determined by the Hamiltonian. We exploit this fact while searching for a Cartan subalgebra. Suppose $\mathfrak{m}$ has the following basis

$$\mathfrak{m} = \text{span}\{\sigma_1, \sigma_2, \ldots, \sigma_{|\mathfrak{m}|}\}, \tag{S.13}$$

where each σ_i is a Pauli string. We construct a list of basis elements for $\mathfrak{h}$ in the following way: we initialize a list of basis elements for $\mathfrak{h}$ via picking a random basis elements from $\mathfrak{m}$, say σ_1. Then we iterate through the basis elements of $\mathfrak{m}$ and append them if they commute with all the Pauli strings we have appended into the list. After a reordering the indices, we obtain the following $\mathfrak{h}$ without loss of generality

$$\mathfrak{h} = \text{span}\{\sigma_1, \sigma_2, \ldots, \sigma_{|\mathfrak{h}|}\}. \tag{S.14}$$

With this notation, $\sigma_i \in \mathfrak{h}$ if $i \leq |\mathfrak{h}|$ and $\sigma_i \notin \mathfrak{h}$ otherwise.

Theorem III.1 *The set given in Eq. (S.14) is a maximal Abelian subalgebra of* $\mathfrak{m}$.

To prove this, we will prove the following lemma:

Lemma III.1 *For* $j = 1, 2, \ldots, |\mathfrak{h}|$, *define the set of non-commuting indices as*

$$s(j) = \{i \,|\, [\sigma_j, \sigma_i] \neq 0\} \tag{S.15}$$

Then

$$\bigcup_{j=1}^{|\mathfrak{h}|} s(j) = \{i \big|\, |\mathfrak{h}| < i \leq |\mathfrak{m}|\} = \{|\mathfrak{h}| + 1, \ldots, |\mathfrak{m}|\}. \tag{S.16}$$

Proof: For any $i, j \leq |\mathfrak{h}|$, we know that $[\sigma_j, \sigma_i] = 0$ because the condition to add the element in $\mathfrak{h}$ is that it commutes with the existing list of basis elements. This implies $i \notin s(j)$ for any $i, j \leq |\mathfrak{h}|$, therefore the union of $s(j)$ sets cannot include any integer smaller than $|\mathfrak{h}| + 1$. For any $i > |\mathfrak{h}|$, we know that we could not add σ_i into $\mathfrak{h}$ during the construction of $\mathfrak{h}$. This means that for any $|\mathfrak{h}| < i \leq |\mathfrak{m}|$, there exists a $j \leq |\mathfrak{h}|$ such that $[\sigma_j, \sigma_i] \neq 0$ i.e. $i \in s(j)$. Therefore union of all $s(1), s(2), \ldots, s(|\mathfrak{h}|)$ must include all positive integers from $|\mathfrak{h}| + 1$ to $|\mathfrak{m}|$. □

Lemma III.2 *For* $j = 1, 2, \ldots, |\mathfrak{h}|$ *and* $i, k = |\mathfrak{h}| + 1, \ldots, |\mathfrak{m}|$, *if* $[\sigma_j, \sigma_i] \neq 0$ *and* $[\sigma_j, \sigma_k] \neq 0$, *then*

$$\text{tr}\left([\sigma_j, \sigma_i][\sigma_j, \sigma_k]\right) = -2^{n+2}\delta_{ik} \tag{S.17}$$

where δ_{ik} *is the Kronecker delta and* σ_i *are* $2^n \times 2^n$ *matrices.*

Proof: Pauli matrices either commute or anti-commute. This extends to Pauli strings as well, because they are tensor products of Pauli matrices. Therefore if $[\sigma_j, \sigma_i] \neq 0$ then $[\sigma_j, \sigma_i] = 2\sigma_j\sigma_i = -2\sigma_i\sigma_j$. We further observe that

$$\begin{aligned} \text{tr}\left([\sigma_j, \sigma_i][\sigma_j, \sigma_k]\right) &= 4\,\text{tr}\left(\sigma_j\sigma_i\sigma_j\sigma_k\right) && \text{(S.18)} \\ &= -4\,\text{tr}\left(\sigma_j\sigma_j\sigma_i\sigma_k\right) && \text{(S.19)} \\ &= -4\,\text{tr}\left(\sigma_i\sigma_k\right) && \text{(S.20)} \end{aligned}$$

Now, if $i \neq k$ i.e. $\sigma_i \neq \sigma_k$, then $\sigma_i\sigma_k$ is not going to be identity: therefore it will be a tensor product of identity and at least one Pauli matrix-matrices. Since $\text{tr}(A \otimes B) = \text{tr}(A)\,\text{tr}(B)$, we then have $\text{tr}(\sigma_i\sigma_k) = 0$ if $i \neq k$. In the $i = k$ case, $\text{tr}(\sigma_i\sigma_k) = \text{tr}(1) = 2^n$. Therefore

$$\text{tr}\left([\sigma_j, \sigma_i][\sigma_j, \sigma_k]\right) = -4 \times 2^n\delta_{ik} = -2^{n+2}\delta_{ik}, \tag{S.21}$$

which proves the lemma. □.

Proof of Theorem III.1: $\mathfrak{h}$ is an Abelian subalgebra of $\mathfrak{m}$ by construction. To prove maximality, we show that there is no other element in $\mathfrak{m}$, including linear combinations of Pauli strings, that commutes with all σ_i for $i \leq |\mathfrak{h}|$. Consider a generic element $m = \sum_{i=1}^{|\mathfrak{m}|} a_i\sigma_i \in \mathfrak{m}$. Then if

$$[\sigma_j, m] = \sum_{i=1}^{|\mathfrak{m}|} a_i[\sigma_j, \sigma_i] = 0 \tag{S.22}$$

for all $j \leq |\mathfrak{h}|$, then we get

$$\sum_{i \in s(j)} a_i [\sigma_j, \sigma_i] = 0 \quad \text{for } j = 1, 2, \ldots, |\mathfrak{h}|. \tag{S.23}$$

Due to orthogonality relations from Lemma III.2, we then obtain $a_i = 0$ for all $i \in \cup_{j \leq |\mathfrak{h}|} s(j)$. Due to Lemma III.1 this implies $a_i = 0$ for all $|\mathfrak{h}| < i \leq |\mathfrak{m}|$, which in turn results in

$$m = \sum_{i=1}^{|\mathfrak{m}|} a_i \sigma_i = \sum_{i=1}^{|\mathfrak{h}|} a_i \sigma_i \in \mathfrak{h}. \tag{S.24}$$

Therefore we have proved that for any $m \in \mathfrak{m}$, if $[\mathfrak{h}, m] = 0$, then $m \in \mathfrak{h}$ which proves maximality. □.

IV. PROOF OF THE THEOREM FOR PARAMETER OPTIMIZATION

In this section, we will point out the differences between Theorem 2 and the methods given in [5, 6]. For this, let us start with stating the original KHK decomposition method as a theorem:

Theorem IV.1 *[5, 6] For $\mathcal{H} \in \mathfrak{m}$, define the function f*

$$f(K) = \langle v, K^\dagger \mathcal{H} K \rangle = \langle K v K^\dagger, \mathcal{H} \rangle, \tag{S.25}$$

where $\langle ., . \rangle$ denotes the Killing form, and $v \in \mathfrak{h}$ is an element whose exponential map e^{itv} is dense in $e^{i\mathfrak{h}}$. Then for any global minimum *of $f(K)$ denoted by K_c,*

$$K_c^\dagger \mathcal{H} K_c \in \mathfrak{h}. \tag{S.26}$$

Although it is stated that a global minimum is needed, both proofs given in [5, 6] only use the fact that the function f has zero gradient at the global minimum, which makes the theorem work for any local extremum as well. By using this fact, we improve the theorem by showing that the element K in the function f does not need to be parameterized via exponential map as in [2, 6, 7] — the parametrization/coordinate system $K = \exp(\sum_i \alpha_i k_i)$ is used to cover the entire Lie group $e^{i\mathfrak{k}}$ which is not generally possible. Because we only need a local extremum, we can re-state the theorem for a generic parameterization system as the following:

Theorem IV.2 ***(Improved KHK Decomposition)*** *Assume a set of coordinates $\vec{\theta}$ in a chart of the Lie group* $\exp(i\mathfrak{k})$*. For $\mathcal{H} \in \mathfrak{m}$, define the function f*

$$f(\vec{\theta}) = \langle K(\vec{\theta}) v K(\vec{\theta})^\dagger, \mathcal{H} \rangle, \tag{S.27}$$

where $\langle ., . \rangle$ denotes an invariant non-degenerate bilinear form on $\mathfrak{g}$, and $v \in \mathfrak{h}$ is an element whose exponential map e^{itv} is dense in $e^{i\mathfrak{h}}$. Then for any local extrema *of $f(\vec{\theta})$ denoted by $\vec{\theta}_c$, and defining the critical group element $K_c = K(\vec{\theta}_c)$, we have*

$$K(\vec{\theta}_c)^\dagger \mathcal{H} K(\vec{\theta}_c) = K_c^\dagger \mathcal{H} K_c \in \mathfrak{h}. \tag{S.28}$$

The motivation for this extension is to use the decoupled product form (Eq. (8) in the main text) rather than the exponential map itself. To prove the theorem, we first provide the following definition.

Definition IV.1 *Given a compact Lie algebra $\mathfrak{k}$ and Lie group $\mathcal{K} = e^{i\mathfrak{k}}$ generated via the exponential map. Let $f : \mathcal{K} \to \mathbb{R}$ be a smooth function. Then, if for a coordinate system $\vec{\theta}$ in $\mathcal{K}$, partial derivatives of f with respect to the coordinates vanish at $K_c \in \mathcal{K}$ and the basis vectors at K_c covers $\mathfrak{k}$ i.e.*

$$\frac{\partial f}{\partial \theta_i}\Big|_{K_c} = 0 \text{ for } i = 1, 2, \ldots, |\mathfrak{k}|, \qquad \text{span}\Big\{ K^\dagger \frac{\partial K}{\partial \theta_i}\Big|_{K_c} \Big\} = \mathfrak{k}, \tag{S.29}$$

then we will denote this critical point K_c as a ***non-singular*** *critical point in the coordinate system $\vec{\theta}$. In the case the basis vectors not covering the $\mathfrak{k}$, we will call it a* ***singular*** *critical point in the coordinate system $\vec{\theta}$.*

The product form (Eq. (8)) might lead to coordinate singularities due to the fact that different set of coordinates might represent the same Lie group element. Therefore, if the product form is used in Theorem IV.2, local extremum might be a singular critical point. We provide the following lemma to cover this case as well.

Lemma IV.1 *If $K_c \in \mathcal{K}$ is a singular critical point of the function $f : \mathcal{K} \to \mathbb{R}$ in the coordinate system $\vec{\theta}$, then there exists a coordinate system $\vec{\alpha}$ such that K_c is a non-singular critical point of f in the coordinates $\vec{\alpha}$.*

Proof: We have

$$\frac{\partial f}{\partial \theta_i}\Big|_{K_c} = 0 \text{ for } i = 1, 2, \ldots, |\mathfrak{k}| \qquad \text{span}\Big\{K^\dagger \frac{\partial K}{\partial \theta_i}\Big|_{K_c}\Big\} \neq \mathfrak{k}. \tag{S.30}$$

Since there are $|\mathfrak{k}|$ many coordinates, some of the basis vectors should be linearly dependent because they fail to span $\mathfrak{k}$. Let us say r of them are linearly independent, where $r < |\mathfrak{k}|$. Without loss of generality, we can choose them to be $\theta_1, \theta_2, ..., \theta_r$. We then conclude that there are r relevant coordinates that we denote as $\vec{\eta}$, and $|\mathfrak{k}| - r$ irrelevant coordinates that we denote as $\vec{\phi}$. That is, denoting the coordinate mapping as $K(\eta, \phi) \in \mathcal{K}$, the Lie group element at the critical point $K_c \in e^{i\mathfrak{k}}$ is determined by the relevant coordinates, and the irrelevant ones do not change the group element: $K_c = K(\vec{\eta}_c, \vec{\phi})$ for any choice of $\vec{\phi}$. In other words, the map from this coordinate system to the Lie group manifold is many-to-one at the critical point K_c. We will exploit the fact that the function f is a function of the group element K. Let us calculate the gradient of f at K_c. Define a smooth curve $C(\lambda)$ on the manifold $C : \mathbb{R}^+ \to \mathcal{K}$ such that the curve passed through at the critical point $C(\lambda = 0) = K_c$. Then gradient of f on the direction of the curve C at $\lambda = 0$ is given by

$$\frac{d}{d\lambda} f(C(\lambda))\Big|_{\lambda=0^+} = \lim_{\lambda \to 0^+} \frac{f(C(\lambda)) - f(K_c)}{\lambda}, \tag{S.31}$$

which is independent of coordinate choice. To write (S.31) in $\vec{\theta}$ coordinates, write $C(\lambda) = K(\vec{\eta}(\lambda), \vec{\phi}(\lambda))$. Then $C(0) = K_c$ implies that $\vec{\eta}(0) = \vec{\eta}_c$ while $\vec{\phi}(0)$ remains free since K_c is independent of $\vec{\phi}$. This leads to

$$\frac{d}{d\lambda} f(C(\lambda))\Big|_{\lambda=0^+} = \lim_{\lambda \to 0^+} \frac{f(K(\vec{\eta}(\lambda), \vec{\phi}(\lambda))) - f(K(\vec{\eta}_c, \vec{\phi}))}{\lambda}. \tag{S.32}$$

Since K_c does not depend on $\vec{\phi}$, we can then write

$$\frac{d}{d\lambda} f(C(\lambda))\Big|_{\lambda=0^+} = \lim_{\lambda \to 0^+} \frac{f(K(\vec{\eta}(\lambda), \vec{\phi}(\lambda))) - f(K(\vec{\eta}_c, \vec{\phi}(\lambda)))}{\lambda} \tag{S.33}$$

$$= \sum_{i=1}^{r} \frac{\partial f}{\partial \eta_i} \frac{d\eta_i}{d\lambda}\Big|_{\lambda=0^+} = \sum_{i=1}^{r} \frac{\partial f}{\partial \theta_i} \frac{d\theta_i}{d\lambda}\Big|_{\lambda=0^+} = 0 \tag{S.34}$$

for any curve C. Let us choose a new set of coordinates

$$\tilde{K}(\vec{\alpha}) = K_c e^{i \sum_i \alpha_i k_i}, \tag{S.35}$$

and rewrite (S.31) with $\vec{\alpha}$ coordinates with $K_c = \tilde{K}(\vec{\alpha} = \vec{0})$ and $C(\lambda) = \tilde{K}(\vec{\alpha}(\lambda))$:

$$0 = \frac{d}{d\lambda} f(C(\lambda))\Big|_{\lambda=0^+} = \lim_{\lambda \to 0^+} \frac{f(\tilde{K}(\vec{\alpha}(\lambda))) - f(\tilde{K}(\vec{0}))}{\lambda} \tag{S.36}$$

$$= \sum_{i=1}^{|\mathfrak{k}|} \frac{\partial f}{\partial \alpha_i} \frac{d\alpha_i}{d\lambda}\Big|_{\lambda=0^+}. \tag{S.37}$$

Since the curve $C(\lambda)$ is arbitrary, $\frac{d\alpha_i}{d\lambda}\big|_{0^+}$ are arbitrary values, which yields

$$\frac{\partial f}{\partial \alpha_i}\Big|_{K_c} = 0, \text{ for } i = 1, 2, \ldots, |\mathfrak{k}|. \tag{S.38}$$

Therefore, K_c is a critical point in the $\vec{\alpha}$ coordinates as well. Now let us show that the basis vectors generated by $\vec{\alpha}$ coordinates will span $\mathfrak{k}$. At $K = K_c$, one can easily see that

$$K^\dagger \frac{\partial K}{\partial \alpha_i}\Big|_{K_c} = K_c^\dagger \lim_{\lambda \to 0} \frac{K_c e^{i\lambda k_i} - K_c}{\lambda} = k_i, \tag{S.39}$$

which yields

$$\text{span}\Big\{K^\dagger \frac{\partial K}{\partial \alpha_i}\Big|_{K_c}\Big\} = \text{span}\Big\{k_i\Big\} = \mathfrak{k} \tag{S.40}$$

Therefore, K_c is a non-singular critical point in the $\vec{\alpha}$ coordinates, which proves the lemma. □

Proof of Theorem IV.2: Let us calculate the partial derivatives of f with respect to θ_i:

$$\begin{aligned}
\frac{\partial f(\vec{\theta})}{\partial \theta_i} &= \Big\langle \frac{\partial K}{\partial \theta_i} v K^\dagger, \mathcal{H} \Big\rangle + \Big\langle K v \frac{\partial K^\dagger}{\partial \theta_i}, \mathcal{H} \Big\rangle \\
&= \Big\langle \frac{\partial K}{\partial \theta_i} v K^\dagger, \mathcal{H} \Big\rangle - \Big\langle K v K^\dagger \frac{\partial K}{\partial \theta_i} K^\dagger, \mathcal{H} \Big\rangle \\
&= \Big\langle K K^\dagger \frac{\partial K}{\partial \theta_i} v K^\dagger, \mathcal{H} \Big\rangle - \Big\langle K v K^\dagger \frac{\partial K}{\partial \theta_i} K^\dagger, \mathcal{H} \Big\rangle \\
&= \Big\langle K \Big[K^\dagger \frac{\partial K}{\partial \theta_i}, v\Big] K^\dagger, \mathcal{H} \Big\rangle
\end{aligned} \tag{S.41}$$

Using similarity transformation invariance of the bilinear form, we can shift $K(...)K^\dagger$ to the right

$$\frac{\partial f(K)}{\partial \theta_i} = \Big\langle \Big[K^\dagger \frac{\partial K}{\partial \theta_i}, v\Big], K^\dagger \mathcal{H} K \Big\rangle \tag{S.42}$$

Now, we can rewrite this as

$$\begin{aligned}
\frac{\partial f(K)}{\partial \theta_i} &= -i \frac{\partial}{\partial t} \Big\langle e^{-itv} K^\dagger \frac{\partial K}{\partial \theta_i} e^{itv}, K^\dagger \mathcal{H} K \Big\rangle \Big|_{t=0} \\
&= -i \frac{\partial}{\partial t} \Big\langle K^\dagger \frac{\partial K}{\partial \theta_i}, e^{itv} K^\dagger \mathcal{H} K e^{-itv} \Big\rangle \Big|_{t=0} \\
&= \Big\langle K^\dagger \frac{\partial K}{\partial \theta_i}, [v, K^\dagger \mathcal{H} K] \Big\rangle
\end{aligned} \tag{S.43}$$

Therefore, at the critical point $K = K_c$

$$\Big\langle K^\dagger \frac{\partial K}{\partial \theta_i}\Big|_{K_c}, [v, K_c^\dagger \mathcal{H} K_c] \Big\rangle = 0. \tag{S.44}$$

We know that $\mathcal{H}, v \in \mathfrak{m}$ and $K_c \in e^{i\mathfrak{k}}$. Therefore by the definition of Cartan decomposition (Def. 1 in the main text), $K_c \mathcal{H} K_c^\dagger \in \mathfrak{m}$ and $[v, K_c^\dagger \mathcal{H} K_c] \in \mathfrak{k}$. (S.43) is satisfied for all $i = 1, 2, \ldots, |\mathfrak{k}|$. By the Lemma IV.1, without loss of generality, we can simply assume that $K^\dagger \frac{\partial K}{\partial \theta_i}\Big|_{K_c}$ span $\mathfrak{k}$. Then due to non-degeneracy of the bilinear form, (S.43) yields

$$[v, K_c^\dagger \mathcal{H} K_c] = 0 \tag{S.45}$$

which also means

$$[e^{itv}, K_c^\dagger \mathcal{H} K_c] = 0. \tag{S.46}$$

The exponential map of v is dense in $e^{i\mathfrak{h}}$, i.e. for any element e^{ih} chosen in the group $e^{i\mathfrak{h}}$, the line e^{itv} passes through a point that is arbitrarily close to the element e^{ih}. This with (S.46) means that $K^\dagger \mathcal{H} K$ commutes with any element in $\mathfrak{h}$. Since $\mathfrak{h}$ is the maximal Abelian Lie algebra in $\mathfrak{m}$, this implies that $K^\dagger \mathcal{H} K \in \mathfrak{h}$. □.

We then conclude that the method given in Refs. [5, 6] **does not** require us to minimize the function f. It requires us to find a **local** extremum. In addition, this gives us flexibility in how to represent the element K in Theorem IV.2. This additional property will be exploited in the next section.

V. PRODUCT ANZATS FOR K

In contrast to other works using Cartan decomposition [2, 6, 7] that represent an element $K \in e^{i\mathfrak{k}}$ as $K = \exp(\sum_i \alpha_i k_i)$, where k_i is the ith basis element of $\mathfrak{k}$, we use the following product representation:

$$K(\vec{\theta}) = \prod_i e^{i\theta_i k_i}, \tag{S.47}$$

For example, for $\mathfrak{k} = \mathfrak{su}(2)$, this product would be

$$K(a,b,c) = e^{iaX} e^{ibY} e^{icZ}. \tag{S.48}$$

The product expansion (S.47) is extremely beneficial when the basis elements k_i are Pauli strings, because it can then be directly implemented in a quantum computer and we only need to find the parameters θ_i to generate the circuit. In [7], Cartan decomposition is applied recursively because the expression $K = \exp(\sum_i \alpha_i k_i)$ cannot be implemented on a quantum device, which is the unitary synthesis problem that we try to solve for Hamiltonian evolution in the first place. A second benefit of this expression comes from the optimization: as explained in the next section, (S.47) allows us to calculate $f(K)$ and its gradient with a greater accuracy.

The product expansion (S.47) does not cover the Lie group $e^{i\mathfrak{k}}$ except special cases such as $\mathfrak{k}$ being Abelian or solvable [8]. However, it is a good parametrization that can be used in the theorem. To show that, let us first notice that the (S.47) is differentiable. Secondly, let us show that (S.47) covers a $|\mathfrak{k}|$ dimensional subspace of $e^{i\mathfrak{k}}$. For this, observe that near identity $K(\vec{\theta} = 0) = I$ (S.47) can be expanded as

$$K(\vec{\theta}) = I + \sum_i \theta_i k_i + O(\theta^2), \tag{S.49}$$

Since each k_i are basis elements of $\mathfrak{k}$, it is obvious that the equation above covers $|\mathfrak{k}|$ dimensional neighborhood of the identity element. Considering the continuity and the differentiability of the parametrization (S.47), we can therefore easily conclude that it covers a $|\mathfrak{k}|$ dimensional subspace of the Lie group $e^{i\mathfrak{k}}$. Therefore this parametrization can be used in Theorem IV.2. Now let us show that the $f(\vec{\theta}) = f(K(\vec{\theta}))$ has a local extremum. For our specific case, we pick our basis elements from single Pauli strings. In this case

$$e^{i\theta_i k_i} = \cos\theta_i I + i \sin\theta_i \, k_i, \tag{S.50}$$

Therefore $f(\vec{\theta})$ is a periodic function on all its variables. Since it is also differentiable, it should have a local extremum within a period. Therefore by using the product expansion, we are guaranteed to find a local extremum and therefore find a solution to our decomposition.

VI. TIME COMPLEXITY FOR PARAMETER OPTIMIZATION

A. Cost function evaluation

To perform the optimization, we need to calculate $f(K) = \langle KvK^\dagger, \mathcal{H} \rangle$ given in Theorem IV.2 (Theorem 2 in the main text) where $\mathcal{H} \in \mathfrak{m}$, $K \in e^{i\mathfrak{k}}$ and v is an element in $\mathfrak{h}$ whose exponential map e^{itv} is dense in $e^{i\mathfrak{h}}$. $\mathfrak{h}$ is an Abelian Lie algebra and in this work, basis elements of $\mathfrak{h}$ are single Pauli strings. Therefore the parameter space for the group $e^{i\mathfrak{h}}$ is 2π periodic on all parameters, meaning that it is a $|\mathfrak{h}|$ dimensional torus. If one chooses v as $a_1 h_1 + a_2 h_2 + ...$ where a_i are mutually irrational to each other, then the line e^{itv} will be dense in $e^{i\mathfrak{h}}$. Therefore we use $v = \sum_i \gamma^i h_i$ where γ is a transcendental number to ensure that any power γ^i is irrational.

We represent K with the following product of exponentials

$$K = \prod_i e^{i\theta_i k_i}, \tag{S.51}$$

where k_i form a basis for $\mathfrak{k}$. Using the fact that Killing form $\langle A, B \rangle$ in $\mathfrak{su}(2^n)$ is proportional to $\mathrm{tr}(AB)$ where tr is the matrix trace, and therefore it is a non-degenerate invariant bilinear form in $\mathfrak{g}(\mathcal{H}) \subset \mathfrak{su}(2^n)$, we can replace $\langle A, B \rangle$ in the function with $\mathrm{tr}(AB)$. Then we find

$$f(K) = \mathrm{tr}\left(\prod_{i\uparrow} e^{i\theta_i k_i} \; v \prod_{i\downarrow} e^{-i\theta_i k_i} \; \mathcal{H} \right), \tag{S.52}$$

where $\uparrow$ ($\downarrow$) under the product means multiplication in an increasing (decreasing) order for i. Efficient calculation of this product is non-trivial because the products of these exponentials generally will not be in $\mathfrak{g}$, and thus may generate an arbitrary matrix in $\mathrm{GL}(2^n)$ and therefore require an exponential amount of calculation. One fact that can be used is that if $K \in e^{i\mathfrak{k}}$ and $m \in \mathfrak{m}$, then we have $KmK^\dagger \in \mathfrak{m}$. Thus, if each exponential on both sides of v in (S.52) is applied on v, one from each side at the same time (a similarity transformation), the result will always be in $\mathfrak{m}$. To apply the exponentials of k_i, one can take advantage of the fact that k_i are Pauli strings, therefore $k_i^2 = I$ and

$$e^{i\theta_i k_i} = \cos\theta_i\, I + i \sin\theta_i\, k_i. \tag{S.53}$$

This allows us to apply one similarity transformation on one term in v via constant amount of calculations only requires a constant amount of calculations. After applying all exponentials to v,

$$f(K) = \mathrm{tr}\left(m_0 \mathcal{H}\right) \tag{S.54}$$

is obtained, where

$$m_0 = \prod_{i\uparrow} e^{i\theta_i k_i}\, v \prod_{i\downarrow} e^{-i\theta_i k_i} \in \mathfrak{m}. \tag{S.55}$$

Up to this point, in the worst case only $\mathcal{O}(|\mathfrak{k}||\mathfrak{m}|)$ many operations are performed. $|\mathfrak{k}|$-many exponentials are applied to an element of $\mathfrak{m}$, which has at most $|\mathfrak{m}|$ many Pauli terms. Multiplying m_0 and $\mathcal{H}$ requires $\mathcal{O}(|\mathfrak{m}|)$ many calculations and can be neglected in the $|\mathfrak{k}| \gg 1$ limit corresponding to large system size limit, which leads to $\mathcal{O}(|\mathfrak{k}||\mathfrak{m}|)$ time complexity to calculate $f(K)$.

B. Gradient evaluation

The gradient of K given in (S.51) can be expressed as

$$\frac{\partial K}{\partial \theta_j} = \prod_{i<j,\uparrow} e^{i\theta_i k_i} i k_j \prod_{i\geq j,\uparrow} e^{i\theta_i k_i}, \tag{S.56}$$

leading to the analytical expression for the gradient of the function (S.52):

$$\begin{aligned}\frac{\partial f(K)}{\partial \theta_j} =& \mathrm{tr}\left(\prod_{i<j,\uparrow} e^{i\theta_i k_i} i k_j \prod_{i\geq j,\uparrow} e^{i\theta_i k_i}\, v \prod_{i\downarrow} e^{-i\theta_i k_i}\, \mathcal{H}\right) \\ &+ \mathrm{tr}\left(\prod_{i\uparrow} e^{i\theta_i k_i}\, v \prod_{i\geq j,\downarrow} e^{i\theta_i k_i}(-i) k_j \prod_{i<j,\downarrow} e^{-i\theta_i k_i}\, \mathcal{H}\right).\end{aligned} \tag{S.57}$$

Using the cyclic property of trace $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ leads to

$$\begin{aligned}\frac{\partial f(K)}{\partial \theta_j} =& i\,\mathrm{tr}\left(k_j \prod_{i\geq j,\uparrow} e^{i\theta_i k_i}\, v \prod_{i\downarrow} e^{-i\theta_i k_i}\, \mathcal{H} \prod_{i<j,\uparrow} e^{i\theta_i k_i}\right) \\ &- i\,\mathrm{tr}\left(\prod_{i<j,\downarrow} e^{-i\theta_i k_i}\, \mathcal{H} \prod_{i\uparrow} e^{i\theta_i k_i}\, v \prod_{i\geq j,\downarrow} e^{i\theta_i k_i}\, k_j\right)\end{aligned} \tag{S.58}$$

Applying the exponentials one by one from both sides of v and $\mathcal{H}$, we obtain

$$\frac{\partial f(K)}{\partial \theta_j} = i\,\mathrm{tr}\left(k_j\, m_1\, e^{-i\theta_j k_j}\, m_2\right) - i\,\mathrm{tr}\left(m_2\, e^{i\theta_j k_j}\, m_1\, k_j\right). \tag{S.59}$$

where

$$\begin{aligned} m_1 &= \prod_{i\geq j,\uparrow} e^{i\theta_i k_i}\, v \prod_{i\geq j\downarrow} e^{-i\theta_i k_i} \in \mathfrak{m} \\ m_2 &= \prod_{i<j\downarrow} e^{-i\theta_i k_i}\, \mathcal{H} \prod_{i<j,\uparrow} e^{i\theta_i k_i} \in \mathfrak{m}. \end{aligned} \tag{S.60}$$

As in the calculation of $f(K)$, reaching that point costs $\mathcal{O}(|\mathfrak{k}||\mathfrak{m}|)$ amount of time and in the $|\mathfrak{k}| \gg 1$ limit corresponding to large system size limit, it is the most time consuming part compared to the last calculation of trace which takes $\mathcal{O}(|\mathfrak{m}|)$ time as above. However, this complexity is to obtain just one derivative. To calculate the full gradient, one has to perform this for all θ_j, and therefore the complexity of calculating the entire gradient is $\mathcal{O}(|\mathfrak{k}|^2|\mathfrak{m}|)$.

C. Obtaining $h \in \mathfrak{h}$

After finding $K_c \in e^{i\mathfrak{k}}$ that locally extremizes $f(K)$, one can obtain $h \in \mathfrak{h}$ via the following (also Eq. (7) in the main text):

$$h = K_c^\dagger \mathcal{H} K_c \in \mathfrak{h}. \tag{S.61}$$

Using the product form of K_c, one finds

$$h = \prod_{i\downarrow} e^{-i\theta_i k_i}\, \mathcal{H} \prod_{i\uparrow} e^{i\theta_i k_i}. \tag{S.62}$$

As discussed in the previous sections, the similarity transformations from each single Pauli exponential $e^{i\theta_i k_i}$ can be applied one by one with analytical precision via (S.53), and since $\mathcal{H} \in \mathfrak{m}$, the terms generated after each similarity transformation are also in $\mathfrak{m}$. This leads to $\mathcal{O}(|\mathfrak{k}||\mathfrak{m}|)$ complexity, just as for the calculation of $f(K)$.

VII. 2 SITE TFIM PARAMETER FIT

As given in Fig. 2, the 2 site transverse field Ising model, $\mathcal{H} = ZZ + B_1 IX + B_2 XI$, has the following Hamiltonian algebra

$$\mathfrak{g}(\mathcal{H}) = \text{span}\{XI, IX, ZZ, YY, YZ, ZY\}, \tag{S.63}$$

and the following Cartan decomposition and Cartan subalgebra are used

$$\begin{aligned} \mathfrak{k} &= \text{span}\{YZ, ZY\}, \\ \mathfrak{m} &= \text{span}\{XI, IX, ZZ, YY\}, \\ \mathfrak{h} &= \text{span}\{XI, IX\}. \end{aligned} \tag{S.64}$$

By defining $v = IX + \gamma XI$, with γ an arbitrary transcendental constant, and $K = e^{iaYZ} e^{ibZY}$, the cost function (7) can be calculated as

$$\begin{aligned} f(a,b) =& \text{tr}\left(e^{iaYZ} e^{ibZY} (IX + \gamma XI) e^{-ibZY} e^{-iaYZ} \mathcal{H}\right) \\ =& (B_1 + \gamma B_2)\cos 2a \cos 2b - (B_2 + \gamma B_1)\sin 2a \sin 2b + \cos 2a \sin 2b + \gamma \sin 2a \cos 2b. \end{aligned} \tag{S.65}$$

To find a local extremum, we set $\partial f/\partial a = \partial f/\partial b = 0$, which yields

$$\begin{aligned} \tan(2a + 2b) &= \frac{1}{B_1 + B_2}, \\ \tan(2a - 2b) &= \frac{1}{B_2 - B_1}, \end{aligned} \tag{S.66}$$

and are solved by

$$\begin{aligned} a &= \frac{1}{4}\arctan\left(\frac{1}{B_1 + B_2}\right) - \frac{1}{4}\arctan\left(\frac{1}{B_1 - B_2}\right), \\ b &= \frac{1}{4}\arctan\left(\frac{1}{B_1 + B_2}\right) + \frac{1}{4}\arctan\left(\frac{1}{B_1 - B_2}\right). \end{aligned} \tag{S.67}$$

Plugging this in $K^\dagger \mathcal{H} K$, one finds

$$\begin{aligned} K^\dagger \mathcal{H} K =& e^{-ibZY} e^{-iaYZ} \mathcal{H} e^{iaYZ} e^{ibZY} \\ =& IX\left(\frac{(B_1 + B_2)^2 - 1}{2\sqrt{1 + (B_1 + B_2)^2}} - \frac{(B_1 - B_2)^2 - 1}{2\sqrt{1 + (B_1 - B_2)^2}}\right) \\ &+ XI\left(\frac{(B_1 + B_2)^2 - 1}{2\sqrt{1 + (B_1 + B_2)^2}} + \frac{(B_1 - B_2)^2 - 1}{2\sqrt{1 + (B_1 - B_2)^2}}\right) \\ =& c\, IX + d\, XI \in \mathfrak{h}. \end{aligned} \tag{S.68}$$

With this, we have $\mathcal{H} = K(cIX + dXI)K^\dagger$, which is the desired relationship.

VIII. CIRCUIT OPTIMIZATION FOR TFXY MODEL

In this section, we outline several circuit optimizations which we apply to the free-fermionizable model discussed in the main text.

Consider an n-qubit circuit. We establish the following Lemma:

Lemma VIII.1 *For any $i, j = 1, 2, ..., n-1$, $i < j$ and any $\alpha, \beta \in \mathbb{R}$, there exist $a, b, c \in \mathbb{R}$ such that*

$$e^{i\alpha \widehat{Y_i X}_j} e^{i\beta \widehat{Y_i X}_{j+1}} = e^{ia\widehat{Y_j X}_{j+1}} e^{ib\widehat{Y_i X}_j} e^{ic\widehat{Y_j X}_{j+1}} \tag{S.69}$$

where the "hat" notation is defined in Eq. 12. The same is true for $X \leftrightarrow Y$.

To prove this, observe that the algebra generated by the exponents of the left hand side is a representation of $\mathfrak{su}(2)$:

$$\begin{aligned} [\widehat{Y_i X}_{j+1}, \widehat{Y_i X}_j] &= 2i\, \widehat{Y_j X}_{j+1} \\ [\widehat{Y_j X}_{j+1}, \widehat{Y_i X}_{j+1}] &= 2i\, \widehat{Y_i X}_j \\ [\widehat{Y_i X}_j, \widehat{Y_j X}_{j+1}] &= 2i\, \widehat{Y_i X}_{j+1} \end{aligned} \tag{S.70}$$

Thus, (S.69) is an Euler decomposition of a $\mathfrak{su}(2)$ spanned by the Pauli strings. The version with $X \leftrightarrow Y$ is also true for the same reason.

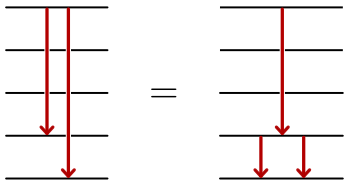

FIG. S1. Graphical representation of Lemma VIII.1 for $i = 1, j = 4$.

Using the arrow notation introduced in Fig. 3, a graphical representation of Lemma VIII.1 can be given as Fig. S1. Next, we apply this Lemma to an ordered product of exponentials as is used in Eqs. (S.52) and (S.56).

Theorem VIII.1 *We define a "triangle" of size i as*

$$T_i(\vec{\alpha}) = \prod_{j=1}^{i} e^{i\alpha_j \widehat{Y_1 X}_j} \tag{S.71}$$

and

$$Z_{p,q}(\vec{\alpha}) = \prod_{j=p}^{q} e^{i\alpha_j \widehat{Y_j X}_{j+1}} \tag{S.72}$$

which will be denoted as "zig" if $p > q$ and "zag" if $p < q$. Then for $i \geq 3$, there exists a new set of parameters $\vec{a}, b, c \in \mathbb{R}$ such that

$$T_i(\vec{\alpha}) = e^{ib\widehat{Y_{j-1} X}_j} T_{i-1}(\vec{a}) e^{ic\widehat{Y_{j-1} X}_j} \tag{S.73}$$

and this implies for a new set of parameters $\vec{\beta}, \vec{\theta} \in \mathbb{R}$, the "triangle" $T_i(\vec{\alpha})$ can be written as a "zigzag"

$$\begin{aligned} T_i(\vec{\alpha}) &= \prod_{j=i-1,\downarrow}^{1} e^{i\beta_j \widehat{Y_j X}_{j+1}} \prod_{j=1}^{i-1} e^{i\theta_j \widehat{Y_j X}_{j+1}} \\ &= Z_{i-1,1}(\vec{\beta}) Z_{2,i-1}(\vec{\theta}). \end{aligned} \tag{S.74}$$

To prove this, we first observe that

$$\begin{aligned} T_i(\vec{\alpha}) &= \prod_{j=1}^{i} e^{i\alpha_j \widehat{Y_1 X_j}} \\ &= \left(\prod_{j=1}^{i-2} e^{i\alpha_j \widehat{Y_1 X_j}} \right) e^{i\alpha_{i-1} \widehat{Y_1 X}_{i-1}} e^{i\alpha_i \widehat{Y_1 X}_i}. \end{aligned} \tag{S.75}$$

Using Lemma VIII.1 on the last two exponentials, and renaming new parameters:

$$\begin{aligned} T_i(\vec{\alpha}) &= \left(\prod_{j=1}^{i-2} e^{ia_j \widehat{Y_1 X_j}} \right) e^{ib\widehat{Y_{i-1} X}_i} e^{ia_{i-1} \widehat{Y_1 X}_{i-1}} e^{ic\widehat{Y_{i-1} X}_i} \\ &= e^{ib\widehat{Y_{i-1} X}_i} \left(\prod_{j=1}^{i-2} e^{ia_j \widehat{Y_1 X_j}} \right) e^{ia_{i-1} \widehat{Y_1 X}_{i-1}} e^{ic\widehat{Y_{i-1} X}_i} \\ &= e^{ib\widehat{Y_{i-1} X}_i} \left(\prod_{j=1}^{i-1} e^{ia_j \widehat{Y_1 X_j}} \right) e^{ic\widehat{Y_{i-1} X}_i} \\ &= e^{ib\widehat{Y_{i-1} X}_i} T_{i-1}(\vec{a}) e^{ic\widehat{Y_{i-1} X}_i}, \end{aligned} \tag{S.76}$$

which is just equation (S.73). Recursively iterating this a total of $i-1$ times, results in Eq. (S.74). In the graphical representation, this recursion is easy to see as shown in Fig. S2.

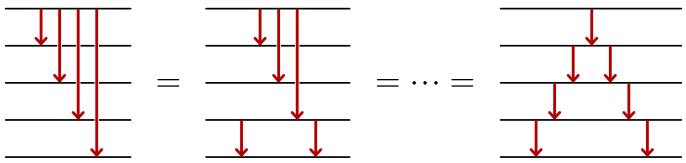

FIG. S2. Graphical representation of Theorem VIII.1 for $i = 5$.

Considering the fact that the original circuit is a product of triangles, we now show that it can be written as a series of zigzags schematically depicted in Fig. S3. This greatly reduces the complexity of the circuit because the initial circuit given on the left has $\mathcal{O}(n^3)$ CNOT gates, whereas the simplified zigzag circuit has only $\mathcal{O}(n^2)$ CNOT gates. However, this circuit can be simplified further.

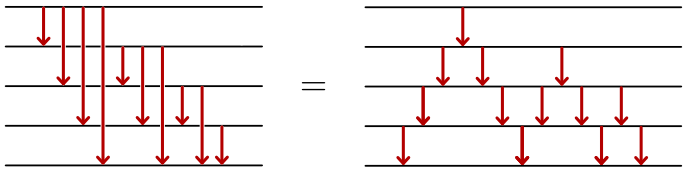

FIG. S3. First simplification of the initial circuit for K.

Lemma VIII.2 *Any zigzag can be flipped into a "zagzig", i.e. for any $i \geq 2$ and any set of parameters $\alpha_j, \beta_j \in \mathbb{R}$, there exists $\vec{a}, \vec{b} \in \mathbb{R}$ such that*

$$Z_{i,1}(\vec{\alpha}) Z_{2,i}(\vec{\beta}) = Z_{1,i}(\vec{a}) Z_{i-1,1}(\vec{b}) \tag{S.77}$$

The proof is by induction. The base case is for zigzags with size $i = 2$, because $\widehat{Y_2 X}_3, \widehat{Y_1 X}_2, \widehat{Y_1 X}_3$ forms a representation of $\mathfrak{su}(2)$ and it is established by using the Euler decomposition in the following two ways:

$$e^{ia\widehat{Y_1 X}_2} e^{ib\widehat{Y_2 X}_3} e^{ic\widehat{Y_1 X}_2} = e^{i\alpha\widehat{Y_2 X}_3} e^{i\beta\widehat{Y_1 X}_2} e^{i\theta\widehat{Y_2 X}_3}. \tag{S.78}$$

This is precisely the $i = 2$ case of the lemma. Now we assume that it also holds for all zigzags up to size N. Then for $i = N + 1$, we have that

$$Z_{N+1,1}(\vec{\alpha})Z_{2,N+1}(\vec{\beta}) = Z_{N+1,3}(\vec{\alpha})\left(e^{i\alpha_2\widehat{Y_2X_3}}e^{i\alpha_1\widehat{Y_1X_2}}e^{i\beta_2\widehat{Y_2X_3}}\right)Z_{3,N+1}(\vec{\beta}) \tag{S.79}$$

The product in parentheses is our base case shown in Eq. (S.78). Therefore, for some $a, b, c \in \mathbb{R}$, we have that

$$\begin{aligned} Z_{N+1,1}(\vec{\alpha})Z_{2,N+1}(\vec{\beta}) =& Z_{N+1,3}(\vec{\alpha})\left(e^{ia\widehat{Y_1X_2}}e^{ib\widehat{Y_2X_3}}e^{ic\widehat{Y_1X_2}}\right)Z_{3,N+1}(\vec{\beta}) \\ =& e^{ia\widehat{Y_1X_2}}\left(Z_{N+1,3}(\vec{\alpha})\ e^{ib\widehat{Y_2X_3}}Z_{3,N+1}(\vec{\beta})\right)e^{ic\widehat{Y_1X_2}}. \end{aligned} \tag{S.80}$$

Note that the expression inside the parentheses on the last line is a zigzag with size N, that runs between sites 2 and $N + 1$. Therefore it can be flipped by the induction hypothesis. After renaming the parameters $a \to a_1$ and $c \to c_1$, we obtain

$$\begin{aligned} Z_{N+1,1}(\vec{\alpha})Z_{2,N+1}(\vec{\beta}) =& e^{ia_1\widehat{Y_1X_2}}\left(Z_{2,N+1}(\vec{a})Z_{N,2}(\vec{b})\right)e^{ic\widehat{Y_1X_2}} \\ =& Z_{1,N+1}(\vec{a})Z_{N,1}(\vec{b}), \end{aligned} \tag{S.81}$$

which proves the induction step. This completes the proof of Lemma VIII.2. A graphical representation is given in Fig. S4.

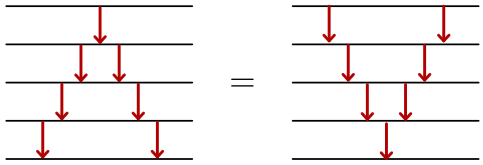

FIG. S4. Graphical representation of Lemma VIII.2 for $i = 5$.

Theorem VIII.2 *The zigzag circuit given for K can be simplified into a more compact multiplications of "zigs", i.e. for every $n \geq m + 1$, $\vec{\alpha}_i, \vec{\beta}_i \in \mathbb{R}$ there exist a set of $\vec{\theta}_i \in \mathbb{R}$, such that*

$$\prod_{i=m}^{n-1}\left(Z_{n,i}(\vec{\alpha}_i)Z_{i+1,n}(\vec{\beta}_i)\right) = \prod_{i=m}^{n} Z_{n,i}(\vec{\theta}_i). \tag{S.82}$$

For convenience, the parameters will not be shown explicitly for this proof i.e. $Z_{n,m}(\vec{\alpha})$ will be written as $Z_{n,m}$, since the parameters are not determined explicitly in the argument. The proof is again by induction. The base case is for $n - m = 1$ since both sides become $Z_{m+1,m}Z_{m,m}$. Now, assume that the induction step holds for all $n - m$ up to $n - m = N \geq 1$. We will next establish that it holds for $n - m = N + 1$. First, define $N' = N + m + 1$ and then regroup the product to obtain

$$\prod_{i=m}^{N'-1}\left(Z_{N',i}Z_{i+1,N'}\right) = Z_{N',m}\prod_{i=m+1}^{N'-1}\left(Z_{i,N'}Z_{N',i}\right)Z_{N',N'}. \tag{S.83}$$

Since $N' > i$, $Z_{i,N'}Z_{N',i}$, the product of terms in the parenthesis can be rewritten as $Z_{i,N'}Z_{N'-1,i}$, yielding

$$= Z_{N',m}\prod_{i=m+1}^{N'-1}\left(Z_{i,N'}Z_{N'-1,i}\right)Z_{N',N'}. \tag{S.84}$$

Using Lemma VIII.2 for the expression in the product, we find that the product becomes

$$\begin{aligned} =& Z_{N',m}\prod_{i=m+1}^{N'-1}\left(Z_{N',i}Z_{i+1,N'}\right)Z_{N',N'} \\ =& Z_{N',m}\prod_{i=m+1}^{N'-1}\left(Z_{N',i}Z_{i,N'}\right). \end{aligned} \tag{S.85}$$

In the last step, we used the fact that the last term in the product in the middle is $Z_{N',N'}$ and therefore can be absorbed into the term after the product by redefining its coefficient in the exponent.

Note that the product term to the right is part of the induction hypothesis for $n-m=N$. Applying the induction hypothesis gives us

$$\begin{aligned} &= Z_{N',m} \prod_{i=m+1}^{N'} \left(Z_{N',i}\right) \\ &= \prod_{i=m}^{N'} \left(Z_{N',i}\right), \end{aligned} \tag{S.86}$$

which proves the induction step, and establishes the theorem. A graphical representation is given in Fig. S5.

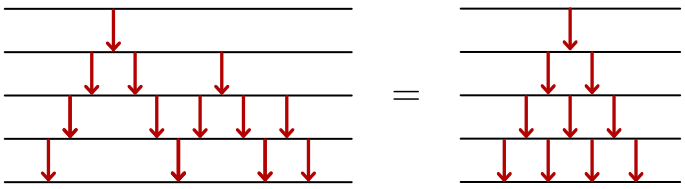

FIG. S5. Graphical representation of theorem VIII.2 for $n=4$ and $m=1$.

Using these results, we find that the red part of the original K circuit, given in the left side of Fig. S3. can be rewritten as the circuit on right shown in Fig. S4. Considering that all the down red arrows commute with all the up green arrows, we can move greens through the reds and arrive at the simplification shown in Fig. S6.

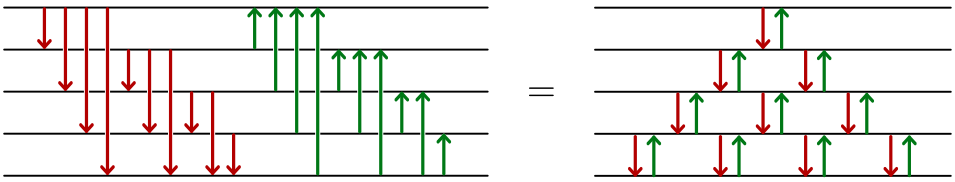

FIG. S6. Simplification of K circuit for Transverse Field XY model for 5 spins.

As shown in Fig. 3(a), an arrow with length a, *i. e.*, the circuit for $\exp(i\theta\widehat{X_iY}_{i+a})$ (or the same circuit, but with $X \leftrightarrow Y$), has $2a$ CNOT gates in it. Therefore, the number of CNOT gates in the circuit on the left of Fig. S6 satisfies

$$\#\text{CNOTs for raw K} = (\text{red part}) + (\text{green part}) = 2\sum_{p=1}^{n-1}\sum_{q=1}^{p}(2q) = 2\sum_{p=1}^{n-1} p(p+1) = \frac{2n(n^2-1)}{3}. \tag{S.87}$$

On the other hand, the optimized circuit for n spins consists pairs of length one red arrows followed by length one green arrows, that is $\exp(i\theta\widehat{Y_iX}_{i+1})\exp(i\phi\widehat{X_iY}_{i+1})$. A circuit for this pair requires only 2 CNOTs [9]. Therefore, the total CNOT count of the simplified circuit on the left of Fig. S6 is reduced to only the following:

$$\#\text{CNOTs for simplified K} = \sum_{p=1}^{n-1}(2p) = n(n-1). \tag{S.88}$$

The full circuit consists of one factor of K, one factor of $\exp(-ith)$ and one factor of $K^\dagger$, as given in Fig. 2(b). Using the Cartan subalgebra given in (11), we see that $\exp(-ith)$ does not require any CNOT gates. Hence, the complete time-evolution circuit of $U(t) = K\exp(-ith)K^\dagger$ has twice as many CNOTs as the circuit for one K has. Therefore, the non-optimized circuit for time evolution has $2n(n^2-1)/3$ CNOTs, whereas the optimized one has only $2n(n-1)$ CNOTs.

[1] M. Dağlı, D. D'Alessandro, and J. Smith, Journal of Physics A Mathematical and Theoretical **41** (2007).

[2] B. Drury and P. Love, Journal of Physics A: Mathematical and Theoretical **41**, 395305 (2008).
[3] D. d'Alessandro, *Introduction to quantum control and dynamics* (CRC press, 2007).
[4] D. D'Alessandro and F. Albertini, Journal of Physics A: Mathematical and Theoretical **40**, 2439 (2007).
[5] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Crm Proceedings & Lecture Notes (American Mathematical Society, 2001) pp. 247–248, Lemma 6.3 (iii).
[6] H. N. Sá Earp and J. K. Pachos, Journal of Mathematical Physics **46**, 082108 (2005), https://doi.org/10.1063/1.2008210.
[7] N. Khaneja and S. J. Glaser, Chemical Physics **267**, 11 (2001).
[8] J. Wei and E. Norman, Proceedings of the American Mathematical Society **15**, 327 (1964).
[9] G. Vidal and C. M. Dawson, Physical Review A **69**, 010301(R) (2004).