Decomposition of Clifford Gates

Tefjol Pllaha, Kalle Volanto, Olav Tirkkonen
Aalto University, Helsinki, Finland
e-mails: {tefjol.pllaha, kalle.volanto, olav.tirkkonen}@aalto.fi

***Abstract*—In fault-tolerant quantum computation and quantum error-correction one is interested on Pauli matrices that commute with a circuit/unitary. This information is encoded by the *support* (Pllaha et al., 2020) of the given circuit/unitary. We provide a fast algorithm that decomposes any Clifford gate as a *minimal* product of Clifford transvections. The algorithm can be directly used for computing the support of any given Clifford gate. To achieve this goal, we exploit the structure of the symplectic group with a novel graphical approach.**

I. Introduction

The Clifford group is of central importance in quantum information and computation. It sits on the second level of the Clifford hierarchy, with the first level being the Heisenberg-Weyl group. This paper is primarily motivated by its importance in fault-tolerant quantum computation and quantum error-correction [1], [2]. Traditionally, the Clifford group is studied via its connection with the binary symplectic group [3] and the associated decompositions of the latter. The Bruhat decomposition of the symplectic group [4] gives a standard generating set made of qubit permutations, diagonal gates, and (partial) Hadamard gates. Alternatively, the Clifford group can be studied via the transvection decomposition of the symplectic group, which we briefly describe in Section III. It is well-known that the symplectic group is generated by symplectic transvections [5], [6]. Although these references give a constructive proof, the decomposition primarily relies on exhaustive search. In this paper we give a simple and fast algorithm that decomposes any symplectic matrix as a *minimal* product of symplectic transvections. This, in turn, gives a decomposition of the Clifford group in terms of *Clifford transvections*. By definition, Clifford transvections are sparse (in fact, they are the most sparse Cliffords other than Paulis and diagonal Cliffords), and given their simple conjugation action, they are also easy to implement. This yields directly a decomposition of any m-qubit Clifford gate as a *minimal* product of Clifford transvections.

In [7], the authors studied the Clifford group via the *support* of a unitary matrix. In that language, Clifford transvections are precisely those Cliffords that have a support of size two, which is smallest support among non-Pauli Cliffords. On top of being a useful algebraic tool, the support of a unitary encodes valuable information about the Paulis that commute with the given unitary. Moreover, the structure of the support can measure the implementation complexity of the gate.

In [7, Prop. 9], the authors compute the support of standard Clifford gates, while leaving the case of a generic Clifford as an open problem and for future research. This is primarily due to the fact that the support behaves unpredictably under multiplication. The results of this paper settle this open problem. Indeed, we provide a fast algorithm, with $\mathcal{O}(m^3)$ worst-case complexity, for computing the support of *any* Clifford gate. Heuristically, we expect our results to have applications in designing flag gadgets [8], [9], [10], [11] for stabilizer circuits, which we will explore in future work. It is also interesting to pursue whether or not similar techniques can be used for the third level of the Clifford hierarchy.

We exploit the structure of symplectic matrices with a novel graphical approach. To a symplectic matrix, written as a minimal product of transvections, one can associate a Gram-type matrix that captures the commutativity relations of the defining transvections along the lines of [12], [13]. When viewed as an adjacency matrix, it defines an *anti-commutation graph*. Unlike [12], [13], we are interested in the *directed* paths of the associated graph, which, as it turns out, completely determine the given symplectic matrix; see Theorem 1. These directed paths can be counted with an invertible upper-triangular matrix, and this allows us to reduce the decomposition problem to a matrix triangulation problem over the binary field. For the latter we make use of the results of [14].

II. Preliminaries

A. The binary symplectic group

The binary symplectic group, denoted $\mathrm{Sp}(2m;2)$, consists of $2m\times 2m$ matrices over the binary field $\mathbb{F}_2$ that preserve the matrix

$$\mathbf{\Omega} = \begin{bmatrix} \mathbf{0}_m & \mathbf{I}_m \\ \mathbf{I}_m & \mathbf{0}_m \end{bmatrix}, \tag{1}$$

under congruence. That is, $\mathbf{F}\in \mathrm{Sp}(2m;2)$ iff $\mathbf{F\Omega F}^{\mathsf{T}} = \mathbf{\Omega}$. Equivalently, symplectic matrices are precisely those matrices that preserve the *symplectic inner product* over $\mathbb{F}_2^{2m}$

$$\langle\, (\mathbf{a},\mathbf{b}) \,|\, (\mathbf{c},\mathbf{d}) \,\rangle_{\mathrm{s}} = \mathbf{ad}^{\mathsf{T}} + \mathbf{bc}^{\mathsf{T}} = (\mathbf{a},\mathbf{b})\mathbf{\Omega}(\mathbf{c},\mathbf{d})^{\mathsf{T}}. \tag{2}$$

We will denote by $\mathrm{GL}(n;2)$ and $\mathrm{Sym}(n;2)$ the groups of $n\times n$ invertible and symmetric matrices over the binary field $\mathbb{F}_2$, respectively. A matrix $\mathbf{F} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \in \mathrm{Sp}(2m;2)$ satisfies $\mathbf{F\Omega F}^{\mathsf{T}} = \mathbf{\Omega}$, which in turn is equivalent with $\mathbf{AB}^{\mathsf{T}}, \mathbf{CD}^{\mathsf{T}} \in \mathrm{Sym}(m;2)$ and $\mathbf{AD}^{\mathsf{T}} + \mathbf{BC}^{\mathsf{T}} = \mathbf{I}_m$. In $\mathrm{Sp}(2m;2)$ we distinguish two subgroups:

$$\mathcal{F}_D := \left\{ \mathbf{F}_D(\mathbf{P}) = \begin{bmatrix} \mathbf{P} & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{P}^{-\mathsf{T}} \end{bmatrix} \,\middle|\, \mathbf{P}\in \mathrm{GL}(m;2) \right\}, \tag{3}$$

$$\mathcal{F}_U := \left\{ \mathbf{F}_U(\mathbf{S}) = \begin{bmatrix} \mathbf{I}_m & \mathbf{S} \\ \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \,\middle|\, \mathbf{S}\in \mathrm{Sym}(m;2) \right\}. \tag{4}$$

Above, $(\bullet)^{-\mathsf{T}}$ denotes the inverse transposed, and directly by definition we have $\mathcal{F}_D \cong \mathrm{GL}(m;2)$ and $\mathcal{F}_U \cong \mathrm{Sym}(m;2)$. Together with matrices

$$\mathbf{F}_{\mathbf{\Omega}}(r) = \begin{bmatrix} \mathbf{I}_{m|-r} & \mathbf{I}_{m|r} \\ \mathbf{I}_{m|r} & \mathbf{I}_{m|-r} \end{bmatrix}, \tag{5}$$

with $\mathbf{I}_{m|r}$ being the block matrix with $\mathbf{I}_r$ in upper left corner and 0 elsewhere, and $\mathbf{I}_{m|-r} = \mathbf{I}_m - \mathbf{I}_{m|r}$, these two groups are the building blocks of the *Bruhat decomposition* with many applications in quantum computation [4], [15]. A symplectic matrix $\mathbf{F} \in \mathrm{Sp}(2m;2)$ is said to be an *involution* if $\mathbf{F}^2 = \mathbf{I}_{2m}$ and is said to be *hyperbolic* if $\langle\, \mathbf{v} \,|\, \mathbf{v}\mathbf{F} \,\rangle_{\mathrm{s}} = 0$ for all $\mathbf{v} \in \mathbb{F}_2^{2m}$. It is straightforward to verify if a hyperbolic map is also an involution. We will denote

$$\mathrm{Fix}(\mathbf{F}) := \ker(\mathbf{I} + \mathbf{F}) := \{\mathbf{v} \in \mathbb{F}_2^{2m} \mid \mathbf{v} = \mathbf{v}\mathbf{F}\}, \tag{6}$$

$$\mathrm{Res}(\mathbf{F}) := \mathrm{rs}\,(\mathbf{I} + \mathbf{F}) := \{\mathbf{v} + \mathbf{v}\mathbf{F} \mid \mathbf{v} \in \mathbb{F}_2^{2m}\}, \tag{7}$$

where $\ker(\bullet)$ and $\mathrm{rs}\,(\bullet)$ denote the null space and the row space of a matrix, respectively. By definition, these spaces satisfy

$$\dim \mathrm{Res}(\mathbf{F}) + \dim \mathrm{Fix}(\mathbf{F}) = 2m. \tag{8}$$

Involutions have the nice property that $\mathrm{Res}(\mathbf{F}) \subseteq \mathrm{Fix}(\mathbf{F})$. Additionally, for an involution we have $\langle\, \mathbf{x} \,|\, \mathbf{y}\mathbf{F} \,\rangle_{\mathrm{s}} = \langle\, \mathbf{x}\mathbf{F} \,|\, \mathbf{y} \,\rangle_{\mathrm{s}}$ and thus $\langle\, \mathbf{x} + \mathbf{x}\mathbf{F} \,|\, \mathbf{y} + \mathbf{y}\mathbf{F} \,\rangle_{\mathrm{s}} = 0$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^{2m}$. This means that $\mathrm{Res}(\mathbf{F})$ is *self-orthogonal* (or *self-dual* if $\dim \mathrm{Res}(\mathbf{F}) = m$) with respect to (2).

B. The Heisenberg-Weyl group

The *bit-flip* and the *phase-flip* gates are given by

$$\mathbf{X} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } \mathbf{Z} := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{9}$$

respectively. For vectors $\mathbf{a}, \mathbf{b} \in \mathbb{F}_2^m$ we will denote

$$\mathbf{D}(\mathbf{a}, \mathbf{b}) := \mathbf{X}^{a_1}\mathbf{Z}^{b_1} \otimes \cdots \otimes \mathbf{X}^{a_m}\mathbf{Z}^{b_m}. \tag{10}$$

The *Heisenberg-Weyl* group is defined as

$$\mathcal{HW}_N := \{i^k \mathbf{D}(\mathbf{a}, \mathbf{b}) \mid \mathbf{a}, \mathbf{b} \in \mathbb{F}_2^m, k \in \mathbb{Z}_4\} \subset \mathbb{U}(N), \tag{11}$$

where $N = 2^m$. We will denote by $\mathcal{PHW}_N := \mathcal{HW}_N / \{\pm\mathbf{I}_N, \pm i\mathbf{I}_N\}$ the *projective* Heisenberg-Weyl group. Hermitian elements of $\mathcal{HW}_N$ are given (and denoted) by $\mathbf{E}(\mathbf{a}, \mathbf{b}) := i^{\mathbf{a}\mathbf{b}^{\mathsf{T}}}\mathbf{D}(\mathbf{a}, \mathbf{b})$.

C. The Clifford group

The *Clifford group* Cliff_N is defined to be the normalizer of $\mathcal{HW}_N$ in $\mathbb{U}(N)$, that is,

$$\mathrm{Cliff}_N := \{\mathbf{G} \in \mathbb{U}(N) \mid \mathbf{G}\,\mathcal{HW}_N\,\mathbf{G}^\dagger \subset \mathcal{HW}_N\}. \tag{12}$$

In order to obtain a finite group, (12) is meant modulo $\mathbb{U}(1)$.

Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_{2m}\}$ be the standard basis of $\mathbb{F}_2^{2m}$, and consider $\mathbf{G} \in \mathrm{Cliff}_N$. Let $\mathbf{c}_i \in \mathbb{F}_2^{2m}$ be such that

$$\mathbf{G}\,\mathbf{E}(\mathbf{e}_i)\,\mathbf{G}^\dagger = \pm\mathbf{E}(\mathbf{c}_i). \tag{13}$$

Then the matrix $\mathbf{F}_{\mathbf{G}}$ whose ith row is $\mathbf{c}_i$ is a symplectic matrix such that

$$\mathbf{G}\,\mathbf{E}(\mathbf{c})\,\mathbf{G}^\dagger = \pm\mathbf{E}(\mathbf{c}\mathbf{F}_{\mathbf{G}}) \tag{14}$$

for all $\mathbf{c} \in \mathbb{F}_2^{2m}$. We thus have a group homomorphism

$$\Phi : \mathrm{Cliff}_N \longrightarrow \mathrm{Sp}(2m;2),\ \mathbf{G} \longmapsto \mathbf{F}_{\mathbf{G}}. \tag{15}$$

In addition, Φ is surjective with kernel $\ker\Phi = \mathcal{PHW}_N$ [16], and thus $\mathrm{Cliff}_N/\mathcal{PHW}_N \cong \mathrm{Sp}(2m;2)$. It follows that Cliff_N is generated by preimages of symplectic matrices (3),(4),(5). Here a preimage $\Phi^{-1}(\mathbf{F})$ is meant up to $\mathcal{HW}_N$. These preimages are, respectively,

$$\mathbf{G}_D(\mathbf{P}) := |\mathbf{v}\rangle \longmapsto |\mathbf{v}\mathbf{P}\rangle, \tag{16}$$

$$\mathbf{G}_U(\mathbf{S}) := \mathrm{diag}\left(i^{\mathbf{v}\mathbf{S}\mathbf{v}^{\mathsf{T}} \mod 4}\right)_{\mathbf{v}\in\mathbb{F}_2^m}, \tag{17}$$

$$\mathbf{G}_{\mathbf{\Omega}}(r) := (\mathbf{H}_2)^{\otimes r} \otimes \mathbf{I}_{2^{m-r}}, \tag{18}$$

where $\mathbf{H}_2$ is the Hadamard gate.

Since Φ is a homomorphism we have that $\Phi(\mathbf{G}^\dagger) = \mathbf{F}_{\mathbf{G}}^{-1}$. It follows that if $\mathbf{G} \in \mathrm{Cliff}_N$ is Hermitian then $\mathbf{F}_{\mathbf{G}}$ is a symplectic involution. Conversely, if $\mathbf{F}$ is a symplectic involution then $\mathbf{G} = \Phi^{-1}(\mathbf{F})$ satisfies $\mathbf{G}^2 \in \mathcal{HW}_N$. As mentioned, a special class of involutions are the hyperbolic maps. If $\mathbf{G} \in \mathrm{Cliff}_N$ corresponds to a hyperbolic $\mathbf{F} \in \mathrm{Sp}(2m;2)$ then (14) implies that $\mathbf{G}\mathbf{E}\mathbf{G}^\dagger$ commutes with $\mathbf{E}$ for all $\mathbf{E}$.

III. Transvection Decomposition of Symplectic Matrices

A *symplectic transvection* is a symplectic map with one-dimensional residue space (7). If $\mathrm{Res}(\mathbf{F}) = \langle \mathbf{v} \rangle$ then the matrix $\mathbf{F} \in \mathrm{Sp}(2m;2)$ must act as

$$\mathbf{T}_{\mathbf{v}} := \mathbf{I} + \mathbf{\Omega}\mathbf{v}^{\mathsf{T}}\mathbf{v},\ \mathbf{x} \longmapsto \mathbf{x} + \langle\, \mathbf{x} \,|\, \mathbf{v} \,\rangle_{\mathrm{s}}\mathbf{v}. \tag{19}$$

We will call two transvections $\mathbf{T}_{\mathbf{v}}, \mathbf{T}_{\mathbf{w}}$ *independent* if the defining $\mathbf{v}, \mathbf{w}$ are independent. Otherwise, we will call the transvections dependent. Note also that $\mathbf{T}_{\mathbf{v}}, \mathbf{T}_{\mathbf{w}}$ *commute*, that is, $\mathbf{T}_{\mathbf{v}} \cdot \mathbf{T}_{\mathbf{w}} = \mathbf{T}_{\mathbf{w}} \cdot \mathbf{T}_{\mathbf{v}}$ iff $\langle\, \mathbf{v} \,|\, \mathbf{w} \,\rangle_{\mathrm{s}} = 0$, that is, iff $\mathbf{v}, \mathbf{w}$ are orthogonal (with respect to (2) of course).

It is well-known that every symplectic $\mathbf{F}$ matrix can be written as a product of symplectic transvections, and the minimal number of transvections needed is closely related with the dimension of the residue space of $\mathbf{F}$. It is shown in [5] that a non-hyperbolic involution $\mathbf{F}$ can be written as a product of r *independent* transvections $\mathbf{T}_{\mathbf{v}_1}, \ldots, \mathbf{T}_{\mathbf{v}_r}$, where $r = r(\mathbf{F}) := \dim \mathrm{Res}(\mathbf{F}) = 2m - \dim \mathrm{Fix}(\mathbf{F})$ and $\mathrm{Res}(\mathbf{F}) = \langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle$. The strategy of [5] is to find $\mathbf{x}$ such that $\langle\, \mathbf{x} \,|\, \mathbf{x}\mathbf{F} \,\rangle_{\mathrm{s}} = 1$ (which exists for non-hyperbolics), and consider $\mathbf{F}\mathbf{T}_{\mathbf{v}}, \mathbf{v} = \mathbf{x} + \mathbf{x}\mathbf{F} \in \mathrm{Res}(\mathbf{F})$, for which $r(\mathbf{F}\mathbf{T}_{\mathbf{v}}) = r(\mathbf{F}) - 1$. One then repeats the process accordingly until a one-dimensional residue space is reached. If $\mathbf{F}$ is a hyperbolic involution, the following lemma shows that it can be written as a product of $r + 1$ transvections, r of which form a basis for $\mathrm{Res}(\mathbf{F})$, and the additional transvection is dependent of the first r.

Lemma 1 ([5, 2.1.8])**.** *Let $\mathbf{F} \in \mathrm{Sp}(2m;2)$ be hyperbolic. Then there exists $\mathbf{v} \in \mathbb{F}_2^{2m}$ such that $\mathbf{F}\mathbf{T}_{\mathbf{v}}$ is non-hyperbolic and $\mathrm{Res}(\mathbf{F}) = \mathrm{Res}(\mathbf{F}\mathbf{T}_{\mathbf{v}})$.*

Proof. Fix any $\mathbf{0} \neq \mathbf{v} = \mathbf{x} + \mathbf{x}\mathbf{F} \in \mathrm{Res}(\mathbf{F})$. Then any $\mathbf{y}$ such that $\langle\, \mathbf{y} \,|\, \mathbf{v} \,\rangle_{\mathrm{s}} = 1 = \langle\, \mathbf{y}\mathbf{F} \,|\, \mathbf{v} \,\rangle_{\mathrm{s}}$ (which of course exists) satisfies $\langle\, \mathbf{y} \,|\, \mathbf{y}\mathbf{F}\mathbf{T}_{\mathbf{v}} \,\rangle_{\mathrm{s}} = 1$, and thus $\mathbf{F}\mathbf{T}_{\mathbf{v}}$ is non-hyperbolic. Next, by

the choice of $\mathbf{v}$, $\mathrm{Res}(\mathbf{F}\mathbf{T}_{\mathbf{v}}) \subseteq \mathrm{Res}(\mathbf{F})$ holds trivially, and equality is due to equal cardinalities. □

This rule breaks, in part, for non-hyperbolic non-involutions. Namely, there are some exceptions, as described in [6], in which such matrices *cannot* be written as a product of r transvections, but require at least $r+1$ transvections. However, these exceptions are rare, and can be treated separately. For this reason, throughout the paper, we will refer to a symplectic matrix as *generic* if it is *not* one of these exceptions.

For involutions (hyperbolic or not) we have the following nicer result.

Proposition 1. *Any involution is a product of commuting transvections. The converse is also true, that is, any product of commuting transvections yields and involution.*

Proof. The result follows immediately by the fact that two transvections commute iff their defining vectors are orthogonal, along with the fact that the residue space of an involution is self-orthogonal. □

A. A Gram-type matrix

Let $\mathbf{F}$ be a symplectic matrix. We associate to a minimal transvection decomposition $\mathbf{F} = \mathbf{T}_{\mathbf{v}_1}\cdots\mathbf{T}_{\mathbf{v}_r}$ a Gram-type matrix

$$\mathbf{A}(\mathbf{v}_1,\ldots,\mathbf{v}_r) \coloneqq [\langle\, \mathbf{v}_i \,|\, \mathbf{v}_j \,\rangle_{\mathrm{s}}]_{i,j} = \mathbf{V}\mathbf{\Omega}\mathbf{V}^{\mathsf{T}}, \tag{20}$$

where $\mathbf{V}$ is the $r \times 2m$ matrix formed by stacking $\mathbf{v}_1,\ldots,\mathbf{v}_r$. Obviously, $\mathbf{A}$ is symmetric and has zero diagonal. Since a minimal transvection decomposition is given by some basis of the residue space, we will assume that $\mathbf{v}_i \in \mathrm{Res}(\mathbf{F})$. Note that $\mathbf{A} = \mathbf{0}$ iff $\mathbf{F}$ is an involution iff $\mathbf{V}$ is self-orthogonal. On the other hand,

$$\begin{aligned}
\langle\, \mathbf{v}_i \,|\, \mathbf{v}_j \,\rangle_{\mathrm{s}} &= \langle\, \mathbf{x}_i + \mathbf{x}_i\mathbf{F} \,|\, \mathbf{x}_j + \mathbf{x}_j\mathbf{F} \,\rangle_{\mathrm{s}} && (21)\\
&= \langle\, \mathbf{x}_i\mathbf{F} \,|\, \mathbf{x}_j \,\rangle_{\mathrm{s}} + \langle\, \mathbf{x}_i \,|\, \mathbf{x}_j\mathbf{F} \,\rangle_{\mathrm{s}} && (22)\\
&= \mathbf{x}_i\mathbf{F}\mathbf{\Omega}\mathbf{x}_j^{\mathsf{T}} + \mathbf{x}_j\mathbf{\Omega}\mathbf{F}^{\mathsf{T}}\mathbf{x}_j^{\mathsf{T}} && (23)\\
&= \mathbf{x}_i(\mathbf{F} + \mathbf{F}^{-1})\mathbf{\Omega}\mathbf{x}_j^{\mathsf{T}} && (24)\\
&= \langle\, \mathbf{x}_i(\mathbf{F} + \mathbf{F}^{-1}) \,|\, \mathbf{x}_j \,\rangle_{\mathrm{s}}. && (25)
\end{aligned}$$

Obviously, $\mathbf{F}$ is an involution iff $\mathbf{F} = \mathbf{F}^{-1}$, and thus $\mathbf{A}$ also captures how far is $\mathbf{F}$ from being an involution, or equivalently, how far is $\mathbf{V}$ from being self-orthogonal.

In what follows we will denote $\mathbf{A}_{\mathrm{u}} \coloneqq \mathtt{triu}(\mathbf{A})$ the upper triangular part of $\mathbf{A}$ and

$$\mathbf{B}(\mathbf{v}_1,\ldots,\mathbf{v}_r) \coloneqq \sum_{\ell=0}^{r-1} \mathbf{A}_{\mathrm{u}}^{\ell}. \tag{26}$$

By definition, $\mathbf{B}$ is upper triangular with all-ones diagonal for any symplectic $\mathbf{F}$, and it is the identity matrix for any involution (since in this case $\mathbf{A} = \mathbf{0}$). Moreover, $\mathbf{A}_{\mathrm{u}}$ is $r \times r$ upper triangular with all-zero diagonal. This yields $\mathbf{A}_{\mathrm{u}}^{r} = \mathbf{0}$, and thus

$$\mathbf{B} = (\mathbf{I}_r + \mathbf{A}_{\mathrm{u}})^{-1} \text{ and } \mathbf{A}_{\mathrm{u}} = \mathbf{I}_r + \mathbf{B}^{-1}. \tag{27}$$

The matrices $\mathbf{A}$ and $\mathbf{B}$ have a natural graphical interpretation. Let us start with $\mathbf{A}$, which can be thought of as the adjacency matrix of the *anti-commutation* graph $\mathcal{G}(\mathbf{V})$ with vertices $\mathbf{v}_i$ and edges $(\mathbf{v}_i, \mathbf{v}_j)$ iff $\langle\, \mathbf{v}_i \,|\, \mathbf{v}_j \,\rangle_{\mathrm{s}} = 1$. On the other hand, its upper triangular part $\mathbf{A}_{\mathrm{u}}$ can be thought of as the adjacency matrix of the corresponding *directed* graph $\mathcal{G}_{\mathrm{d}}(\mathbf{V})$ with edges $(\mathbf{v}_i, \mathbf{v}_j)$ iff $\langle\, \mathbf{v}_i \,|\, \mathbf{v}_j \,\rangle_{\mathrm{s}} = 1$ *and* $i < j$. As for the matrix $\mathbf{B}$, note first that entry (i,j) of $\mathbf{A}_{\mathrm{u}}^{\ell}$ counts directed paths from $\mathbf{v}_i$ to $\mathbf{v}_j$ of length ℓ. Thus, entry (i,j) (always for $i < j$) counts the number of directed paths from $\mathbf{v}_i$ to $\mathbf{v}_j$.

Remark 1. The matrix $\mathbf{A}(\mathbf{V})$ and the associated anti-commutation graph $\mathcal{G}(\mathbf{V})$ have been discussed in [12], [13], where the authors focus on the orbits of the action of $\{\mathbf{T}_{\mathbf{v}} \mid \mathbf{v} \in \mathbf{V}\}$ on $\mathbb{F}_2^{2m}$. As we will see, for our purposes, $\mathcal{G}(\mathbf{V})$ encodes only limited information, and we will have to make use of $\mathcal{G}_d(\mathbf{V})$ and the matrix $\mathbf{B}(\mathbf{V})$.

Remark 2. The matrix $\mathbf{A}(\mathbf{V})$ is a binary matrix whereas the matrix $\mathbf{B}(\mathbf{V})$ is an integer matrix. However, all the equations that involve $\mathbf{B}(\mathbf{V})$ (e.g., (27) or (28)) are over $\mathbb{F}_2$.

Before providing an example of the notions introduced, we point out that the matrix $\mathbf{B}$ also captures the number of *distinct* transvection decompositions of a given symplectic matrix $\mathbf{F}$. However, this treatment goes beyond the scope of this paper and will be presented in future work.

Example 1. Let us consider an example with $m = 5$ and $\mathbf{F} = \mathbf{T}_{\mathbf{v}_1}\mathbf{T}_{\mathbf{v}_2}\mathbf{T}_{\mathbf{v}_3}\mathbf{T}_{\mathbf{v}_4}\mathbf{T}_{\mathbf{v}_5}$, where

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \\ \mathbf{v}_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then one computes

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 & 3 & 4 \\ 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The graphical description of this scenario is given in Figure 1. For instance, entry $b_{1,4} = 3$ and there are precisely three directed paths from $\mathbf{v}_1$ to $\mathbf{v}_4$, namely, $(\mathbf{v}_1, \mathbf{v}_4)$, $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4)$, and $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$.

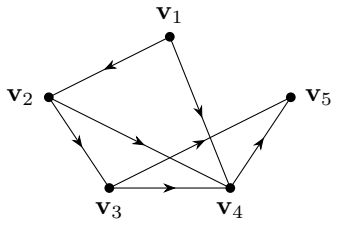

Fig. 1. The directed graph with adjacency matrix $\mathbf{A}_{\mathrm{u}}$.

Theorem 1. *For any symplectic matrix* $\mathbf{F} = \mathbf{T}_{\mathbf{v}_1}\cdots\mathbf{T}_{\mathbf{v}_r}$ *we have*

$$\mathbf{F} = \mathbf{I} + \mathbf{\Omega}\mathbf{V}^{\mathsf{T}}\mathbf{B}\mathbf{V}. \tag{28}$$

As a consequence, if $\mathbf{F}$ *is an involution then* $\mathbf{F} = \mathbf{I} + \mathbf{\Omega}\mathbf{V}^{\mathsf{T}}\mathbf{V}$

Proof. By the definition of transvections, the action of $\mathbf{F}$ on $\mathbf{x}$ is given by some linear combination of $\mathbf{v}_j$ added to $\mathbf{x}$, that is

$$\mathbf{x}\mathbf{F} = \mathbf{x} + \sum_{j=1}^{r} w_j \mathbf{v}_j, \tag{29}$$

where w_j depends on $\langle \mathbf{x} | \mathbf{v}_i \rangle_{\mathrm{s}}$ for $i < j$. We claim that $w_j = \mathbf{x}\mathbf{\Omega}\mathbf{V}^{\mathsf{T}}\mathbf{B}_j$ where $\mathbf{B}_j$ is the jth column of $\mathbf{B}$. This in turn will complete the proof. In order to prove the claim, note that the input of $\mathbf{T}_{\mathbf{v}_j}$ is $\mathbf{x}\mathbf{T}_{\mathbf{v}_1}\cdots\mathbf{T}_{\mathbf{v}_{j-1}}$. Thus $\langle \mathbf{x} | \mathbf{v}_i \rangle_{\mathrm{s}}$ contributes to w_j only if $i < j$ *and* there is a directed path from $\mathbf{v}_i$ to $\mathbf{v}_j$, which could be of length $1 \le \ell \le j - i$. This information is precisely encoded by $\mathbf{B}_j$. □

To the best of our knowledge, Theorem 1 constitutes a novel structural result about symplectic matrices, and comparing it with (19), should come as no surprise. This structure is the main building block of what follows. Based on Theorem 1, it is imperative to consider the *residue matrix*

$$\widehat{\mathbf{F}} := \mathbf{\Omega}(\mathbf{I} + \mathbf{F}) = \mathbf{V}^{\mathsf{T}}\mathbf{B}\mathbf{V} = \sum_{i,j} b_{i,j}\mathbf{v}_i^{\mathsf{T}}\mathbf{v}_j. \tag{30}$$

The terminology comes from the obvious fact that $\mathrm{rs}(\widehat{\mathbf{F}}) = \mathrm{Res}(\mathbf{F})$. Note that $\widehat{\mathbf{F}}$ is symmetric iff $\mathbf{B} = \mathbf{I}$ (recall that $\mathbf{B}$ is upper triangular) iff $\mathbf{F}$ is an involution. Moreover, since $\widehat{\mathbf{F}^{\mathsf{T}}}$ has all-zero diagonal iff $\widehat{\mathbf{F}}$ does, and since

$$\mathbf{x}\widehat{\mathbf{F}^{\mathsf{T}}}\mathbf{x}^{\mathsf{T}} = \mathbf{x}\mathbf{\Omega}(\mathbf{I} + \mathbf{F}^{\mathsf{T}})\mathbf{x}^{\mathsf{T}} = \mathbf{x}\mathbf{\Omega}\mathbf{x}^{\mathsf{T}} + \mathbf{x}\mathbf{\Omega}\mathbf{F}^{\mathsf{T}}\mathbf{x}^{\mathsf{T}} = \langle \mathbf{x} | \mathbf{x}\mathbf{F} \rangle_{\mathrm{s}}, \tag{31}$$

we conclude that $\widehat{\mathbf{F}}$ has all-zero diagonal iff $\mathbf{F}$ is hyperbolic. In such case $\mathbf{F}$ is also an involution, and thus $\widehat{\mathbf{F}}$ is *alternating* (that is, symmetric and all-zero diagonal). It follows by Lemma 1 that we may restrict ourselves to non-hyperbolic maps, and thus we will assume that $\widehat{\mathbf{F}}$ *is not* alternating.

B. Decomposition of Symplectic Involutions

In this subsection we will present a simple algorithm for the decomposition of (non-hyperbolic) symplectic involutions, which will also provide intuition for the much more delicate decomposition of generic symplectic matrices.

Theorem 2 (Transvection Decomposition of Involutions)**.** *Let $\mathbf{F}$ be a non-hyperbolic involution, so that the residue matrix $\widehat{\mathbf{F}}$ is non-alternating. Then there exists $\mathbf{P} \in \mathrm{GL}(2m; 2)$ such that $\mathbf{F} = \mathbf{T}_{\mathbf{v}_1}\cdots\mathbf{T}_{\mathbf{v}_r}$, where $r = \dim \mathrm{Res}(\mathbf{F})$ and $\mathbf{v}_j$ is the jth row of $\mathbf{P}\widehat{\mathbf{F}}$ for $1 \le j \le r$.*

Proof. Let $\mathbf{R}$ be the matrix of row operations that transforms $\widehat{\mathbf{F}}$ into Row-Reduced Echelon form. Let $\mathbf{E}$ be the $r \times r$ upper left block of $\mathbf{R}\widehat{\mathbf{F}}\mathbf{R}^{\mathsf{T}}$, which is invertible by construction. It will also be symmetric and have non-zero diagonal since $\mathbf{F}$ is non-hyperbolic involution. Then there exists $\mathbf{Q} \in \mathrm{GL}(r; 2)$ such that $\mathbf{Q}\mathbf{E}\mathbf{Q}^{\mathsf{T}} = \mathbf{I}_r$; see [5, 2.1.14] for instance. Now put $\mathbf{P} = \mathtt{blkdiag}(\mathbf{Q}, \mathbf{I}_{2m-r})\mathbf{R}$. Then

$$\mathbf{P}\widehat{\mathbf{F}}\mathbf{P}^{\mathsf{T}} = \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{32}$$

We will consider the nonzero rows of $\mathbf{P}\widehat{\mathbf{F}}$, that is, $\begin{bmatrix} \mathbf{Q}\mathbf{E} & \mathbf{0} \end{bmatrix}\mathbf{R}^{-\mathsf{T}}$. For $1 \le j \le r$ let $\mathbf{w}_j$ denote the jth row of $\mathbf{P}\widehat{\mathbf{F}}$, that is, $\mathbf{w}_j = \mathbf{e}_j\mathbf{P}^{-\mathsf{T}}$, where $\mathbf{e}_j \in \mathbb{F}_2^{2m}$ is the jth standard basis vector. Put $\mathbf{F}' = \mathbf{T}_{\mathbf{w}_1}\cdots\mathbf{T}_{\mathbf{w}_r}$. Since $\mathbf{w}_j$'s are linear combinations of $\mathbf{v}_j$'s and since $\mathbf{F}$ is an involution it follows that $\mathbf{A}(\mathbf{w}_1, \ldots, \mathbf{w}_r) = \mathbf{A}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \mathbf{0}_r$ and $\mathbf{B}(\mathbf{w}_1, \ldots, \mathbf{w}_r) = \mathbf{B}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \mathbf{I}_r$. Then (30) yields

$$\widehat{\mathbf{F}'} = \sum_{j=1}^{r} \mathbf{w}_j^{\mathsf{T}}\mathbf{w}_j = \sum_{j=1}^{r} \mathbf{P}^{-1}\mathbf{e}_j^{\mathsf{T}}\mathbf{e}_j\mathbf{P}^{-\mathsf{T}} = \mathbf{P}^{-1}\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{P}^{-\mathsf{T}} = \widehat{\mathbf{F}}, \tag{33}$$

and thus $\mathbf{F} = \mathbf{F}'$. □

The strength of Theorem 2 is that, as we will see, it can be generalized to non-involutions. The case of involutions can be dealt separately with an alternate approach, which, however, does not generalize to non-involutions. According to [17, Thm. 4.1], an involution $\mathbf{F}$ is conjugate with an involution of form

$$\mathbf{F}_U(\mathbf{S}) \equiv \begin{bmatrix} \mathbf{I} & \mathbf{S} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{S} \in \mathrm{Sym}(m; 2), \tag{34}$$

that is, there exists $\mathbf{M} \in \mathrm{Sp}(2m; 2)$ such that $\mathbf{M}\mathbf{F}\mathbf{M}^{-1} = \mathbf{F}_U(\mathbf{S})$. On the other hand, the involutions of form (34) are easy to decompose as described in [7, Prop. 9(2)]. So let us assume $\mathbf{F}_U(\mathbf{S}) = \mathbf{T}_{\mathbf{v}_1}\cdots\mathbf{T}_{\mathbf{v}_r}$. It is straightforward to verify that $\mathbf{T}_{\mathbf{v}}\mathbf{M} = \mathbf{M}\mathbf{T}_{\mathbf{v}\mathbf{M}}$ holds for any symplectic $\mathbf{M}$. This yields

$$\mathbf{F} = \mathbf{M}^{-1}\mathbf{F}_U(\mathbf{S})\mathbf{M} \tag{35}$$

$$= (\mathbf{M}^{-1}\mathbf{T}_{\mathbf{v}_1}\mathbf{M}) \cdot (\mathbf{M}^{-1}\mathbf{T}_{\mathbf{v}_2}\mathbf{M})\cdots(\mathbf{M}^{-1}\mathbf{T}_{\mathbf{v}_r}\mathbf{M}) \tag{36}$$

$$= \mathbf{T}_{\mathbf{v}_1\mathbf{M}}\cdots\mathbf{T}_{\mathbf{v}_r\mathbf{M}}. \tag{37}$$

C. Decomposition of Symplectic Matrices

Finding a transvection decomposition for involutions is facilitated by the simple nature of their associated $\mathbf{A}$ and $\mathbf{B}$ matrices. As we will see, the general case is much more complicated. Let $\mathbf{F}$ be a generic non-hyperbolic symplectic matrix and consider its residue matrix $\widehat{\mathbf{F}}$, for which $\mathrm{rank}(\widehat{\mathbf{F}}) = r$. Thus, a transvection decomposition of $\mathbf{F}$ is given by some basis of $\mathrm{Res}(\mathbf{F}) = \mathrm{rs}(\widehat{\mathbf{F}})$. The task in hand is how to find such basis. The main idea is to start with some fixed basis and transform it accordingly until we reach the desired result. We will start with a basis of $\mathrm{Res}(\mathbf{F})$ in Row-Reduced Echelon form, that is, let $\mathbf{R}$ be a matrix of row operations so that $\mathbf{R}\widehat{\mathbf{F}} = \begin{bmatrix} \mathbf{V} \\ \mathbf{0} \end{bmatrix}$, where $\mathbf{V}$ is a $r \times 2m$ basis. This can be done, for instance, via Gauss Elimination over $\mathbb{F}_2$. Then

$$\mathbf{R}\widehat{\mathbf{F}}\mathbf{R}^{\mathsf{T}} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{E} \in \mathrm{GL}(r; 2). \tag{38}$$

As mentioned, the basis $\mathbf{V}$ may or may not constitute a transvection decomposition of $\mathbf{F}$, and the idea is to consider other bases of form $\mathbf{Q}\mathbf{V}$ where $\mathbf{Q} \in \mathrm{GL}(r; 2)$. Let us denote $\mathbf{P} = \mathtt{blkdiag}(\mathbf{Q}, \mathbf{I}_{2m-r})$, and let $\mathbf{B} = \mathbf{B}(\mathbf{Q}\mathbf{V})$.

Lemma 2. *With the same notation as above, the basis $\mathbf{Q}\mathbf{V}$ constitutes a transvection decomposition of $\mathbf{F}$ iff $\mathbf{Q}\mathbf{E}\mathbf{Q}^{\mathsf{T}} = \mathbf{B}^{-\mathsf{T}}$.*

Proof. Assume $\mathbf{QV}$ gives a transvection decomposition for $\mathbf{F}$. Then, by (30) we have $\widehat{\mathbf{F}} = (\mathbf{QV})^{\mathsf{T}} \cdot \mathbf{B} \cdot (\mathbf{QV})$. But with the notation above we have $\mathbf{QV} = \begin{bmatrix} \mathbf{QE} & \mathbf{0} \end{bmatrix} \mathbf{R}^{-\mathsf{T}}$. Thus

$$\widehat{\mathbf{F}} = \mathbf{V}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}} \cdot \mathbf{B} \cdot \mathbf{QV} \tag{39}$$

$$= \mathbf{R}^{-1} \begin{bmatrix} \mathbf{E}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}} \\ \mathbf{0} \end{bmatrix} \cdot \mathbf{B} \cdot \begin{bmatrix} \mathbf{QE} & \mathbf{0} \end{bmatrix} \mathbf{R}^{-\mathsf{T}} \tag{40}$$

$$= \mathbf{R}^{-1} \begin{bmatrix} \mathbf{E}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}} \cdot \mathbf{B} \cdot \mathbf{QE} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{R}^{-\mathsf{T}}. \tag{41}$$

It follows by (38) that $\mathbf{E} = \mathbf{E}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}} \cdot \mathbf{B} \cdot \mathbf{QE}$ and thus $\mathbf{QEQ}^{\mathsf{T}} = \mathbf{B}^{-\mathsf{T}}$. The reverse direction follows similarly. □

Lemma 3. *With the same notation as above, if* $\mathbf{QEQ}^{\mathsf{T}}$ *is lower triangular, then* $\mathbf{QEQ}^{\mathsf{T}} = \mathbf{B}^{-\mathsf{T}}$.

Proof. Assume $\mathbf{QEQ}^{\mathsf{T}} = \mathbf{L}$ is lower triangular. It follows that

$$\mathbf{L} = (\mathbf{QE}) \cdot \mathbf{E}^{-\mathsf{T}} \cdot (\mathbf{QE})^{\mathsf{T}} \tag{42}$$

and

$$\mathbf{E} = (\mathbf{QE})^{\mathsf{T}} \cdot \mathbf{L}^{-\mathsf{T}} \cdot (\mathbf{QE}). \tag{43}$$

Then, by (38) we have

$$\widehat{\mathbf{F}} = \mathbf{R}^{-1} \begin{bmatrix} (\mathbf{QE})^{\mathsf{T}} \cdot \mathbf{L}^{-\mathsf{T}} \cdot (\mathbf{QE}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{R}^{-\mathsf{T}} \tag{44}$$

$$= \mathbf{R}^{-1} \begin{bmatrix} \mathbf{E}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}} \\ \mathbf{0} \end{bmatrix} \cdot \mathbf{L}^{-\mathsf{T}} \cdot \begin{bmatrix} \mathbf{QE} & \mathbf{0} \end{bmatrix} \mathbf{R}^{-\mathsf{T}} \tag{45}$$

$$= (\mathbf{QV})^{\mathsf{T}} \cdot \mathbf{L}^{-\mathsf{T}} \cdot (\mathbf{QV}). \tag{46}$$

Since $\mathbf{F} = \mathbf{I} + \mathbf{\Omega}\widehat{\mathbf{F}}$ is symplectic, (46) implies

$$\mathbf{\Omega} = \mathbf{F}^{\mathsf{T}}\mathbf{\Omega}\mathbf{F} \tag{47}$$

$$= \mathbf{\Omega} + (\mathbf{QV})^{\mathsf{T}} \cdot (\mathbf{L}^{-\mathsf{T}} + \mathbf{L}^{-1} + \mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-\mathsf{T}}) \cdot (\mathbf{QV}), \tag{48}$$

where $\mathbf{A} = \mathbf{A}(\mathbf{QV}) = (\mathbf{QV})\mathbf{\Omega}(\mathbf{QV})^{\mathsf{T}}$ as defined in (20). Therefore

$$(\mathbf{QV})^{\mathsf{T}} \cdot (\mathbf{L}^{-\mathsf{T}} + \mathbf{L}^{-1} + \mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-\mathsf{T}}) \cdot (\mathbf{QV}) = \mathbf{0}, \tag{49}$$

and since the rows of $\mathbf{QV}$ are linearly independent,

$$\mathbf{L}^{-\mathsf{T}} + \mathbf{L}^{-1} + \mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-\mathsf{T}} = \mathbf{0}. \tag{50}$$

Multiplying from the left by $\mathbf{L}$ and from the right by $\mathbf{L}^{\mathsf{T}}$ gives

$$\mathbf{L} + \mathbf{L}^{\mathsf{T}} = \mathbf{A} = \mathbf{A}_{\mathrm{l}} + \mathbf{A}_{\mathrm{u}}, \tag{51}$$

where $\mathbf{A}_{\mathrm{l}} = \mathbf{A}_{\mathrm{u}}^{\mathsf{T}}$ denotes the lower triangular part of the symmetric matrix $\mathbf{A}$. Using (27), along with the fact that $\mathbf{L}$ is invertible and lower triangular, we obtain

$$\mathbf{L} = \mathbf{I}_r + \mathbf{A}_{\mathrm{l}} = \mathbf{B}^{-\mathsf{T}}, \tag{52}$$

which in turn concludes the proof. □

It follows by Lemmas 2 and 3 that we are seeking for matrices $\mathbf{Q}$ that triangularize $\mathbf{E}$ from (38) by congruence. For more on triangularizations by congruence and related algorithms we refer the reader to [14]. The exceptional non-hyperbolic non-involutions (that require an additional transvection) can be understood in terms of non-triangulable matrices, which we will present in future work. It also follows by Lemma 2 that $\widehat{\mathbf{F}}$ *can* be triangularized by congruence for any non-hyperbolic $\mathbf{F}$ (since for this, one would only need a transvection decomposition of $\mathbf{F}$, which we know it always exists). We resume everything to the following theorem.

Theorem 3 (Transvection Decomposition of Symplectic Matrices)**.** *Let* $\mathbf{F}$ *be a* generic *symplectic matrix. Then there exists an algorithm that for any generic symplectic matrix* $\mathbf{F}$ *outputs a minimal transvection decomposition.*

Proof. If the residue matrix $\widehat{\mathbf{F}}$ is alternating, that is, if $\mathbf{F}$ is hyperbolic, then pick $\mathbf{v}$ as in Lemma 1 and update the input $\mathbf{F}$ with the non-hyperbolic $\mathbf{F}\mathbf{T}_{\mathbf{v}}$, while keeping the residue space intact. Next, perform Gauss Elimination on $\widehat{\mathbf{F}}$ with $\mathbf{R}$ as in (38), and let $\mathbf{Q}$ be such that $\mathbf{QEQ}^{\mathsf{T}}$ is lower triangular. Then, by Lemmas 2 and 3, the r nonzero rows of $\mathtt{blkdiag}(\mathbf{Q}, \mathbf{I}_{2m-r})\mathbf{R}\widehat{\mathbf{F}}$, where $r = \dim \mathrm{Res}(\mathbf{F})$, along with $\mathbf{v}$, yield a minimal transvection decomposition for $\mathbf{F}$. □

IV. Decomposition of Clifford Gates

In [7], the authors studied the Clifford hierarchy via the support of the underlying gates. Every gate $\mathbf{U} \in \mathbb{U}(N)$ can be written as

$$\mathbf{U} = \frac{1}{N} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} \mathrm{Tr}\big(\mathbf{E}(\mathbf{v})\mathbf{U}\big)\, \mathbf{E}(\mathbf{v}), \tag{53}$$

and the support of $\mathbf{U}$ consist of the basis terms that appear in (53), that is,

$$\mathrm{supp}(\mathbf{U}) \coloneqq \{\mathbf{E}(\mathbf{v}) \in \mathcal{HW}_N \mid \mathrm{Tr}\big(\mathbf{E}(\mathbf{v})\mathbf{U}\big) \neq 0\}. \tag{54}$$

On the other hand, (15) assigns $\mathbf{F} \in \mathrm{Sp}(2m; 2)$ to a coset $\mathcal{HW}_N\mathbf{G} = \Phi^{-1}(\mathbf{F})$ for any $\mathbf{G} \in \mathrm{Cliff}_N$. The Clifford

$$\mathbf{G}_{\mathbf{v}} \coloneqq \frac{\mathbf{I}_N \pm i\mathbf{E}(\mathbf{v})}{\sqrt{2}} \in \mathrm{Cliff}_N \tag{55}$$

corresponds to the transvection $\mathbf{T}_{\mathbf{v}}$. Then, since every symplectic matrix is a product of transvections, it follows that

$$\mathbf{G} = \mathbf{E}_0 \prod_{n=1}^{k} \frac{\mathbf{I}_N + i\mathbf{E}_n}{\sqrt{2}} = \frac{\mathbf{E}_0}{\sqrt{|S|}} \sum_{\mathbf{E} \in S} \alpha_{\mathbf{E}}\mathbf{E}, \tag{56}$$

where $\mathbf{E}_0 \in \mathcal{HW}_N, S = \langle \mathbf{E}_1, \ldots, \mathbf{E}_k \rangle$, and $\alpha_{\mathbf{E}} \in \{\pm 1, \pm i\}$; see [7, Prop. 4]. We see from (56) that the support of a Clifford matrix has the additional feature that it is a subgroup[1] of $\mathcal{HW}_N$, which, given the isomorphism $\mathbf{E}(\mathbf{v}) \longleftrightarrow \mathbf{v}$, can be equivalently thought of as a subspace of $\mathbb{F}_2^{2m}$. From earlier discussion, it follows that the support of any $\mathbf{G} \in \Phi^{-1}(\mathbf{F})$ is given by $\mathrm{Res}(\mathbf{F})$ if $\mathbf{F}$ is non-hyperbolic, and by some subspace of $\mathrm{Res}(\mathbf{F})$ of index 2 otherwise. In [7, Prop. 9], the authors determined the support of the standard Clifford gates (16)-(18), while the general case remained open. The difficulty arose by the fact that the support of products is hard to compute. This problem can now be solved with the aid of Theorem 3, as resumed in Algorithm 1. The worst-case complexity of the algorithm is $\mathcal{O}(m^3)$ stemming from the triangularization process over $\mathbb{F}_2$. It is also worth mentioning that in this process one may lose an eighth root of unity; see

[1] To be precise, the support of a Clifford $\mathbf{G}$, written as in (56), is a subgroup iff $\mathbf{E}_0 \in \mathrm{supp}(\mathbf{G})$ iff $\mathrm{Tr}(\mathbf{G}) \neq 0$, and otherwise it is a coset.

Algorithm 1 Transvection Decomposition of Clifford Gates

Input: A Clifford gate $\mathbf{G}$.

1. Compute $\mathbf{F}$ from (13).
2. Compute $\mathbf{v}, \mathbf{v}_1, \cdots, \mathbf{v}_r$ from Theorem 3.
3. $\mathbf{G}_0 = \mathbf{G}_\mathbf{v} \prod_j \mathbf{G}_{\mathbf{v}_j}$.
4. Find $\mathbf{E}_0 = \mathbf{E}(\mathbf{v}_0)$ such that $\mathbf{G} = \mathbf{E}_0\mathbf{G}_0$.

Output: $\mathbf{v}_0$, $\mathbf{v}_j$'s.

Example 3 for instance. We point out here that $\mathbf{G}$ is traceless iff $\mathbf{E}_0 \notin S$. Thus the search in Step 4. of Algorithm 1 can be reduced to either outside S if $\mathbf{G}$ is traceless or in S otherwise.

Example 2. The Hadamard gate can be written as

$$\mathbf{H}_2 = \frac{1}{\sqrt{2}}(\mathbf{X}+\mathbf{Z}) = \mathbf{X}\frac{\mathbf{I}+i\mathbf{Y}}{\sqrt{2}}, \tag{57}$$

where $\mathbf{Y} = i\mathbf{X}\mathbf{Z}$ as usual. Consider now the m fold transversal Hadamard gate $\mathbf{H}_N = (\mathbf{H}_2)^{\otimes m}$, for which $\Phi(\mathbf{H}_N) = \mathbf{\Omega}$. Additionally $\widehat{\mathbf{\Omega}} = \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{bmatrix}$ and $\dim \mathrm{Res}(\mathbf{\Omega}) = m$. Then $\mathbf{P} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{I} \end{bmatrix}$ triangularizes $\widehat{\mathbf{\Omega}}$:

$$\mathbf{P}\widehat{\mathbf{\Omega}}\mathbf{P}^\mathsf{T} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{58}$$

The first m nonzero rows of $\mathbf{P}\widehat{\mathbf{\Omega}}$ are $\begin{bmatrix} \mathbf{I} & \mathbf{I} \end{bmatrix}$. We see that the nth row yields the gate $\mathbf{Y}_n$ with $\mathbf{Y}$ in qubit n and identity elsewhere. From Step 3. of Algorithm 1 we compute

$$\mathbf{G}_0 = \prod_{n=1}^{m} \frac{\mathbf{I}_N + i\mathbf{Y}_n}{\sqrt{2}}. \tag{59}$$

We then find $\mathbf{H}_N = \mathbf{X}^{\otimes m}\mathbf{G}_0$. A similar result holds for partial Hadamard gates $\mathbf{H}^{\otimes r} \otimes \mathbf{I}_{2^{m-r}}$, to which correspond symplectics of form (5); see also [7, Prop. 9(3)]

Example 3. The symplectic and residue matrices corresponding to the CNOT gate are given by

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \text{ and } \widehat{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{60}$$

from which we see that $\widehat{\mathbf{F}}$ is alternating, and thus $\mathbf{F}$ is hyperbolic. So first, we transform $\mathbf{F}$ to a non-hyperbolic map by using the first non-zero row of $\widehat{\mathbf{F}}$, that is, $\mathbf{v} = 0010$. Then we update $\mathbf{F} \leftarrow \mathbf{F}\mathbf{T}_\mathbf{v}$, for which

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \text{ and } \widehat{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{61}$$

A matrix that triangularizes $\widehat{\mathbf{F}}$ is given by

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{62}$$

The non-zero rows of $\mathbf{P}\widehat{\mathbf{F}}$ are $\mathbf{v}_1 = 0100, \mathbf{v}_2 = 0110$. Note that $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, and $\langle \mathbf{v}_1 \,|\, \mathbf{v}_2 \rangle_\mathrm{s} = 0$ as in Proposition 1. Then we compute

$$\mathbf{G}_0 = \frac{(\mathbf{I}+i\mathbf{I}\otimes\mathbf{X})(\mathbf{I}-i\mathbf{Z}\otimes\mathbf{X})(\mathbf{I}+i\mathbf{Z}\otimes\mathbf{I})}{\sqrt{8}}. \tag{63}$$

And then we end with the observation that CNOT= $\xi\mathbf{G}_0$, where $\xi = (1-i)/\sqrt{2}$ is an eighth root of unity.

ACKNOWLEDGEMENTS

This work was funded in part by the Academy of Finland (grant 334539).

REFERENCES

[1] D. Gottesman, "Stabilizer codes and quantum error correction," *PhD thesis, California Institute of Technology*, 1997.

[2] R. A. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction via codes over $\mathrm{GF}(4)$," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1369–1387, 1998.

[3] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction and orthogonal geometry," *Phys. Rev. Lett.*, vol. 78, no. 3, pp. 405–408, 1997. [Online]. Available: https://doi.org/10.1103/PhysRevLett.78.405

[4] D. Maslov and M. Roetteler, "Shorter stabilizer circuits via Bruhat decomposition and quantum circuit transformations," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 4729–4738, 2018.

[5] O. T. O'Meara, *Symplectic groups*, ser. Mathematical Surveys. American Mathematical Society, Providence, R.I., 1978, vol. 16.

[6] D. Callan, "The generation of $\mathrm{Sp}(\mathbb{F}_2)$ by transvections," *J. Algebra*, vol. 42, no. 2, pp. 378–390, 1976.

[7] T. Pllaha, N. Rengaswamy, O. Tirkkonen, and R. Calderbank, "Un-Weyl-ing the Clifford Hierarchy," *Quantum*, vol. 4, p. 370, Dec. 2020. [Online]. Available: https://doi.org/10.22331/q-2020-12-11-370

[8] C. Chamberland and M. E. Beverland, "Flag fault-tolerant error correction with arbitrary distance codes," *Quantum*, vol. 2, p. 53, Feb. 2018. [Online]. Available: https://doi.org/10.22331/q-2018-02-08-53

[9] C. Chamberland and A. W. Cross, "Fault-tolerant magic state preparation with flag qubits," *Quantum*, vol. 3, p. 143, May 2019. [Online]. Available: https://doi.org/10.22331/q-2019-05-20-143

[10] R. Chao and B. W. Reichardt, "Quantum Error Correction with Only Two Extra Qubits," *Phys. Rev. Lett.*, vol. 121, no. 5, p. 050502, 2018.

[11] ——, "Flag fault-tolerant error correction for any stabilizer code," *PRX Quantum*, vol. 1, p. 010302, Sep 2020.

[12] R. Brown and S. P. Humphries, "Orbits Under Symplectic Transvections I," *Proceedings of the London Mathematical Society*, vol. s3-52, no. 3, pp. 517–531, 05 1986. [Online]. Available: https://doi.org/10.1112/plms/s3-52.3.517

[13] ——, "Orbits Under Symplectic Transvections II: The Case $K = \mathbb{F}_2$," *Proceedings of the London Mathematical Society*, vol. s3-52, no. 3, pp. 532–556, 1986. [Online]. Available: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s3-52.3.532

[14] J. D. Botha, "Triangularizing matrices over GF(2) by congruence," *Linear and Multilinear Algebra*, vol. 42, no. 2, pp. 109–158, 1997. [Online]. Available: https://doi.org/10.1080/03081089708818495

[15] T. Pllaha, O. Tirkkonen, and R. A. Calderbank, "Reconstruction of multi-user binary subspace chirps," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 531–536.

[16] N. Rengaswamy, R. A. Calderbank, H. D. Pfister, and S. Kadhe, "Synthesis of logical clifford operators via symplectic geometry," in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 791–795.

[17] S. Beigi and P. W. Shor, "$\mathcal{C}_3$, Semi-Clifford and Generalized Semi-Clifford Operations," *Quantum Inf. Comput.*, vol. 10, pp. 0041–0059, 2010. [Online]. Available: http://arxiv.org/abs/0810.5108