

# Approximate Unitary 3-Designs from Transvection Markov Chains

Xinyu Tan · Narayanan Rengaswamy · Robert Calderbank

the date of receipt and acceptance should be inserted later

**Abstract** Unitary  $k$ -designs are probabilistic ensembles of unitary matrices whose first  $k$  statistical moments match that of the full unitary group endowed with the Haar measure. In prior work, we showed that the automorphism group of classical  $\mathbb{Z}_4$ -linear Kerdock codes maps to a unitary 2-design, which established a new classical-quantum connection via graph states. In this paper, we construct a Markov process that mixes this Kerdock 2-design with symplectic transvections, and show that this process produces an  $\epsilon$ -approximate unitary 3-design. We construct a graph whose vertices are Pauli matrices, and two vertices are connected by directed edges if and only if they commute. A unitary ensemble that is transitive on vertices, edges, and non-edges of this Pauli graph is an exact 3-design, and the stationary distribution of our process possesses this property. With respect to the symmetries of Kerdock codes, the Pauli graph has two types of edges; the Kerdock 2-design mixes edges of the same type, and the transvections mix the types. More precisely, on  $m$  qubits, the process samples  $O(\log(N^5/\epsilon))$  random transvections, where  $N = 2^m$ , followed by a random Kerdock 2-design element and a random Pauli matrix. Hence, the simplicity of the protocol might make it attractive for several applications. From a hardware perspective, 2-qubit transvections exactly map to the Mølmer-Sørensen gates that form the native 2-qubit operations for trapped-ion quantum computers. Thus, it might be possible to extend our work to construct an approximate 3-design that only involves such 2-qubit transvections.

**Keywords** Pauli group, Markov chains, Clifford group, symplectic transvections, unitary designs

## 1 Introduction

Unitary  $k$ -designs are finite collections of  $m$ -qubit unitary matrices endowed with a probability distribution, and they satisfy a certain statistical criterion. Set  $N := 2^m$ . Then, the first  $k$  statistical moments of this finite ensemble match that of the unique rotationally-invariant Haar measure  $\eta(\cdot)$  on the group  $\mathbb{U}_N$  of all  $N \times N$  unitary operators  $\rho$ , i.e.,

$$\sum_i \alpha_i U_i^{\otimes k} \rho (U_i^\dagger)^{\otimes k} = \int_{\mathbb{U}_N} d\eta(U) U^{\otimes k} \rho (U^\dagger)^{\otimes k} \quad (1)$$

for ensemble  $\mathcal{E} = \{\alpha_i, U_i\}_{i=1}^n$  where the unitary matrix  $U_i$  is selected with probability  $\alpha_i$ .

Unitary designs serve several purposes in quantum information science such as quantum data hiding [1], decoupling in quantum information theory [2–5], and channel fidelity estimation [6].

---

X. Tan · R. Calderbank

Department of Mathematics, Duke University, Durham, North Carolina 27708, USA.

N. Rengaswamy

Department of Electrical and Computer Engineering, The University of Arizona, Tucson, Arizona 85721, USA. Most of this work was conducted when he was with the Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708, USA.

E-mail: {xinyu.tan, robert.calderbank}@duke.edu, narayananr@arizona.edu

*Randomized benchmarking* is a protocol commonly used to estimate the quality of state preparation, measurement and operations (gates) in a quantum computer [7, 8]. The actual errors in the system could be gate- and time-dependent, so estimating the error environment completely is challenging. Therefore, the procedure attempts to estimate the *average gate fidelity* that characterizes the error environment. The protocol works by first preparing a fixed initial state  $|\psi\rangle$ . Then, for a fixed sequence length  $s$ , one generates  $K_s$  sequences of  $(s+1)$  operations each, where the first  $s$  operations are chosen randomly from a unitary 2-design and the last operation is defined to be the inverse of the composition of the first  $s$  operations. Hence, ideally, the final state should be identical to the initial state. For each sequence, one measures the average survival probability of the state, i.e.,  $\text{Tr}[\Pi_\psi \mathcal{S}(|\psi\rangle\langle\psi|)]$ , where  $\Pi_\psi$  is the POVM element to detect  $|\psi\rangle$  and  $\mathcal{S}(\cdot)$  is the effective channel induced by the aforementioned sequence of (noisy) operations. If the POVM is realized ideally, then  $\Pi_\psi = |\psi\rangle\langle\psi|$ . Then, one averages the results over the  $K_s$  sequences to compute the average sequence fidelity. This procedure is repeated for each  $s$  and the results are fit to a fidelity decay function.

This scheme amounts to “twirling” the underlying noise channel using a unitary 2-design in order to arrive at the depolarizing channel with the same fidelity as the original channel [7, 8]. Then, the fidelity is estimated on this single-parameter depolarizing channel in order to establish the quality of the computing environment. Small, practically feasible unitary 2-designs make it possible to efficiently characterize the reliability of a quantum computing environment, thereby enabling the development of quantum computers.

In prior work, we showed that the unitary 2-design constructed by Cleve et al. [9] coincided with the symmetry group of the  $\mathbb{Z}_4$ -linear Kerdock code [10]. The Kerdock codewords appear as graph states, providing a new connection between classical and quantum information theory. Our 2-design is a subgroup of the Clifford group, and the corresponding group of binary symplectic matrices,  $\mathfrak{P}_m$ , is isomorphic to the projective special linear group  $\text{PSL}(2, 2^m)$ .

It is well-known that the Clifford group is an exact unitary 3-design, meaning no proper subgroup of the Clifford group gives a 3-design [11]. In this paper, we combine our (Kerdock) 2-design with symplectic transvections [12–14] to construct a Markov process that results in an approximate unitary 3-design. Hence, our work demonstrates how one can “smoothly” turn the Kerdock 2-design into a 3-design.

Symplectic transvections play the same role in the symplectic group as the reflections play in the orthogonal group. More formally, the symplectic transvection  $\tau_h$  corresponding to the row vector  $h \in \mathbb{F}_2^{2m}$  is the map  $\tau_h: \mathbb{F}_2^{2m} \rightarrow \mathbb{F}_2^{2m}$  defined by

$$\tau_h(x) := x + \langle x, h \rangle_s h, \quad (2)$$

where  $\langle \cdot, \cdot \rangle_s$  denotes the symplectic inner product. More details can be found in Section 2.4.

**Main Result (Theorem 5):** Random sampling of  $O(\log(N^5/\epsilon))$  transvections, followed by a random element from  $\mathfrak{P}_m \cong \text{PSL}(2, 2^m)$  and a random Pauli matrix, produces an  $\epsilon$ -approximate unitary 3-design.

An exact unitary 3-design must be transitive on all Paulis, on ordered pairs of commuting Paulis, and on ordered pairs of anti-commuting Paulis. The Kerdock 2-design acts transitively on Pauli elements, but partitions the Pauli pairs into multiple orbits. Using a finite field representation of Paulis, and the fact that transvections generate the symplectic group, we characterize the orbits of Pauli pairs and analyze how each transvection acts on the orbits. Finally, we analyze the convergence rate to an  $\epsilon$ -approximate 3-design using the second largest eigenvalues of the transition matrices of the “edge”- and “orbit”-Markov chains.

While only few applications exist currently for unitary 3-designs [15–17], we think that the simplicity of our protocol makes it an attractive candidate for any such application. Our Markov process samples Clifford transformations uniformly from cosets of the Kerdock 2-design that are determined by products of transvections. Since transvections form a conjugacy class inside the Clifford group (see (21)), the intermediate Cliffords in the Markov process (from  $\mathfrak{P}_m \cong \text{PSL}(2, 2^m)$ ) can be combined into one final Clifford (from  $\mathfrak{P}_m \cong \text{PSL}(2, 2^m)$ ) as stated in the result above. We emphasize here that, while the full Clifford group forms a 3-design [11], transvections form a specific subset whose structure could be exploited for practical implementations. In particular, 2-qubit transvections exactly correspond to Mølmer-Sørensen gates that form the native 2-qubit operations in a trapped-ion quantum computer [18, 19]. Therefore, it might be possible to suitably modify our approach in this paper to construct an approximate 3-design with only 2-qubit transvections. This way, the design could be tailor-made for trapped-ion systems.

The rest of the paper is organized as follows. In Section 2, we introduce our finite field representation of the Pauli group elements, show how it expresses commutativity, and discuss other preliminaries including the Clifford group and symplectic transvections. In Section 3, we define a graph on Pauli matrices, where Clifford elements act as graph automorphisms, and explain how the Pauli graph works with the ideas of Pauli mixing and Pauli 2-mixing. In Section 4, we introduce the Kerdock unitary 2-design which is a symplectic subgroup isomorphic to the projective special linear group  $\text{PSL}(2, 2^m)$ . Then, using our finite field representation, we define orbit invariants based on how  $\text{PSL}(2, 2^m)$  partitions the directed edges of the Pauli graph. In Section 5, we introduce the transvection Markov chains whose stationary distribution gives an exact unitary 3-design. In Section 6, we analyze the convergence rate of the transvection Markov process and prove that it produces an  $\epsilon$ -approximate unitary 3-design. In Section 7, we conclude the paper.

## 2 The Pauli and Clifford Groups

In this section, we describe commutativity in the Pauli group by rewriting the standard symplectic inner product as a trace inner product over a finite field. This translation simplifies the description of the Clifford symmetries that generate the Kerdock unitary 2-design.

### 2.1 The Finite Field $\mathbb{F}_{2^m}$

The field representation is fundamental to our description of commutativity in the Pauli group. We obtain the finite field  $\mathbb{F}_{2^m}$  from the binary field  $\mathbb{F}_2$  by adjoining a root  $\alpha$  of a primitive irreducible polynomial  $p(x)$  of degree  $m$  [20]. Each element of  $\mathbb{F}_{2^m}$  corresponds to a polynomial in  $\alpha$  of degree at most  $m - 1$  with coefficients in  $\mathbb{F}_2$ . The field elements  $1, \alpha, \alpha^2, \dots, \alpha^{m-1}$  form a basis for  $\mathbb{F}_{2^m}$  over  $\mathbb{F}_2$  which we call the *primal basis*. The corresponding *dual basis*  $\beta_0, \beta_1, \dots, \beta_{m-1}$  is defined by

$$\text{Tr}(\alpha^i \beta_j) := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \quad (3)$$

where the *trace*  $\text{Tr}: \mathbb{F}_{2^m} \rightarrow \mathbb{F}_2$  is the  $\mathbb{F}_2$  linear map

$$\text{Tr}(x) := x + x^2 + \dots + x^{2^{m-1}}. \quad (4)$$

Given a field element  $a \in \mathbb{F}_{2^m}$ , we will write

$$a = \sum_{i=0}^{m-1} [a]_i \alpha^i = \sum_{i=0}^{m-1} [a]_i \beta_i. \quad (5)$$

The binary row vector  $[a] \in \{0, 1\}^m$  represents the coefficients of  $a$  in the primal basis, and the binary row vector  $[a] \in \{0, 1\}^m$  represents the coefficients of  $a$  in the dual basis.

The trace is linear over  $\mathbb{F}_2$  and the binary symmetric matrix  $W$  given by

$$W_{ij} := \text{Tr}(\alpha^i \alpha^j), \quad i, j = 0, 1, \dots, m-1 \quad (6)$$

satisfies

$$[a] = [a]W, \quad (7)$$

thereby translating primal coordinates to dual coordinates.

The trace inner product  $\text{Tr}(ab) = \langle a, b \rangle_{\text{tr}}$  is given by

$$\text{Tr}(ab) = [a]W[b]^T = [a] \cdot [b] \pmod{2}. \quad (8)$$

The matrix  $W$  is non-singular since the trace inner product is non-degenerate (if  $\text{Tr}(xz) = 0$  for all  $z \in \mathbb{F}_{2^m}$  then  $x = 0$ ). Observe that  $W$  is a Hankel matrix, since if  $i + j = h + k$  then  $\text{Tr}(\alpha^i \alpha^j) = \text{Tr}(\alpha^h \alpha^k)$ .

## 2.2 The Pauli Group

The single qubit (Hermitian) *Pauli* matrices are

$$I_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad Y := iXZ = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad (9)$$

where  $i := \sqrt{-1}$  and  $I_2$  is the  $2 \times 2$  identity matrix [21].

Pauli matrices on  $m$  qubits are described by Kronecker products of  $m$  single qubit Pauli matrices. We associate with each pair of finite field elements  $(a, b) \in \mathbb{F}_{2^m} \times \mathbb{F}_{2^m}$  the  $m$ -fold Kronecker product

$$D(a, b) := X^{[a]_0} Z^{[b]_0} \otimes \dots \otimes X^{[a]_{m-1}} Z^{[b]_{m-1}} \in \mathbb{U}_N, \quad (10)$$

where  $N := 2^m$  and  $\mathbb{U}_N$  denotes the group of all  $N \times N$  unitary operators.

The  $m$ -qubit *Pauli group*  $P_N$  (also called the *Heisenberg-Weyl group*) consists of all operators  $i^\kappa D(a, b)$ . The values  $i^\kappa$ , where  $\kappa \in \mathbb{Z}_4 := \{0, 1, 2, 3\}$ , are called *quaternary phases*. The order  $|P_N| = 4N^2$  and the *center* of this group is  $\langle iI_N \rangle := \{I_N, iI_N, -I_N, -iI_N\}$ , where  $I_N$  is the  $N \times N$  identity matrix. Hence, the homomorphism  $\psi: P_N \rightarrow \mathbb{F}_2^{2m}$  defined by

$$\psi(i^\kappa D(a, b)) := [[a], [b]] \quad \forall \kappa \in \mathbb{Z}_4 \quad (11)$$

has kernel  $\langle iI_N \rangle$  and allows us to represent elements of  $P_N$  (up to multiplication by scalars) as binary row vectors or pairs of  $\mathbb{F}_{2^m}$  elements.

Multiplication in  $P_N$  satisfies the identity

$$D(a, b)D(c, d) = (-1)^{[c] \cdot [b] + [a] \cdot [d]} D(c, d)D(a, b). \quad (12)$$

The standard *symplectic inner product* in  $\mathbb{F}_2^{2m}$  is defined as

$$\begin{aligned} \langle [[a], [b]], [[c], [d]] \rangle_s &:= [c] \cdot [b] + [a] \cdot [d] \\ &= [[a], [b]] \Omega [[c], [d]]^\top, \end{aligned} \quad (13)$$

where the symplectic form  $\Omega := \begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix}$  (see [14, 22, 23]). It follows from (8) that

$$\langle [[a], [b]], [[c], [d]] \rangle_s = \text{Tr}(ad + bc). \quad (14)$$

Therefore, two operators  $D(a, b)$  and  $D(c, d)$  commute if and only if  $\text{Tr}(ad + bc) = 0$ .

## 2.3 The Clifford Group

The *Clifford group*  $\text{Cliff}_N$  is the *normalizer* of  $P_N$  in the unitary group  $\mathbb{U}_N$ . It consists of all unitary matrices  $g \in \mathbb{C}^{N \times N}$  for which  $gD(a, b)g^\dagger \in P_N$  for all  $D(a, b) \in P_N$ , where  $g^\dagger$  is the Hermitian transpose of  $g$  [24].

The Clifford group contains  $P_N$  and has size  $|\text{Cliff}_N| = 2^{m^2+2m} \prod_{j=1}^m (4^j - 1)$  (ignoring scalars  $e^{2\pi i\theta}$ ,  $\theta \in \mathbb{R}$ ) [22]. Every operator  $g \in \text{Cliff}_N$  induces an automorphism of  $P_N$  by conjugation. Note that the inner automorphisms induced by matrices in  $P_N$  preserve every conjugacy class  $\{\pm D(a, b)\}$  and  $\{\pm iD(a, b)\}$ , because (12) implies that elements in  $P_N$  either commute or anti-commute. Matrices  $D(a, b)$  are symmetric or anti-symmetric according as  $\text{Tr}(ab) = 0$  or 1, hence the matrix

$$E(a, b) := i^{\text{Tr}(ab)} D(a, b) \quad (15)$$

is Hermitian. Note that  $E(a, b)^2 = I_N$ .

The automorphism induced by a Clifford element  $g$  satisfies

$$gE(a, b)g^\dagger = \pm E([[a], [b]]F_g), \quad \text{where } F_g = \begin{bmatrix} A_g & B_g \\ C_g & D_g \end{bmatrix} \quad (16)$$

**Table 1** A generating set of symplectic matrices and their corresponding unitary operators.

The number of 1s in  $Q$  and  $P$  directly relates to number of gates involved in the circuit realizing the respective unitary operators (see [14, 23, Appendix I]). The  $N$  coordinates are indexed by binary vectors  $v \in \mathbb{F}_2^m$ , and  $e_v$  denotes the standard basis vector in  $\mathbb{C}^N$  with an entry 1 in position  $v$  and all other entries 0. Here  $H_{2^t}$  denotes the Walsh-Hadamard matrix of size  $2^t$ ,  $U_t = \text{diag}(I_t, 0_{m-t})$  and  $L_{m-t} = \text{diag}(0_t, I_{m-t})$ .

Symplectic Matrix $F_g$	Clifford Operator $g$
$\Omega = \begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix}$	$H_N = H_2^{\otimes m}$
$L_Q = \begin{bmatrix} Q & 0 \\ 0 & Q^{-\top} \end{bmatrix}$	$\ell_Q : e_v \mapsto e_{vQ}$
$T_P = \begin{bmatrix} I_m & P \\ 0 & I_m \end{bmatrix}; P = P^\top$	$t_P = \text{diag}(i^{vPv^\top \bmod 4})$
$G_t = \begin{bmatrix} L_{m-t} & U_t \\ U_t & L_{m-t} \end{bmatrix}$	$g_t = H_{2^t} \otimes I_{2^{m-t}}$

is a  $2m \times 2m$  binary matrix that preserves symplectic inner products:

$$\langle [[a], [b]], [[c], [d]] \rangle_{F_g} = \langle [[a], [b]], [[c], [d]] \rangle_s. \quad (17)$$

Hence,  $F_g$  is called a *binary symplectic matrix* and the symplectic property reduces to  $F_g \Omega F_g^\top = \Omega$ , or equivalently

$$A_g B_g^\top = B_g A_g^\top, \quad C_g D_g^\top = D_g C_g^\top, \quad A_g D_g^\top + B_g C_g^\top = I_m. \quad (18)$$

The symplectic property encodes the fact that the automorphism induced by  $g$  must respect commutativity in  $P_N$ . Let  $\text{Sp}(2m, \mathbb{F}_2)$  denote the group of symplectic  $2m \times 2m$  matrices over  $\mathbb{F}_2$ . The map  $\phi: \text{Cliff}_N \rightarrow \text{Sp}(2m, \mathbb{F}_2)$  defined by

$$\phi(g) := F_g \quad (19)$$

is a homomorphism with kernel  $P_N$ , and every Clifford operator projects onto a symplectic matrix  $F_g$ . Thus,  $P_N$  is a normal subgroup of  $\text{Cliff}_N$  and  $\text{Cliff}_N/P_N \cong \text{Sp}(2m, \mathbb{F}_2)$ . This implies that  $|\text{Sp}(2m, \mathbb{F}_2)| = 2^{m^2} \prod_{j=1}^m (4^j - 1)$  (also see [22]).

Table 1 lists elementary symplectic transformations  $F_g$ , that generate the binary symplectic group  $\text{Sp}(2m, \mathbb{F}_2)$ , and the corresponding unitary automorphisms  $g \in \text{Cliff}_N$ , which together with  $P_N$  generate  $\text{Cliff}_N$ . (See [23, Appendix I] for a discussion on the Clifford gates and circuits corresponding to these transformations.)

## 2.4 Symplectic Transvections

The symplectic transvection  $\tau_h$  corresponding to the row vector  $h \in \mathbb{F}_2^{2m}$  is the map  $\tau_h: \mathbb{F}_2^{2m} \rightarrow \mathbb{F}_2^{2m}$  defined by

$$\tau_h(x) := x + \langle x, h \rangle_s h = x Z_h, \quad (20)$$

where  $Z_h := I_{2m} + \Omega h^\top h$  is the associated symplectic matrix of  $\tau_h$ .

The set of symplectic transvections is a conjugacy class in  $\text{Sp}(2m, \mathbb{F}_2)$ , since if  $h = [h_1, h_2]$  with  $h_1, h_2 \in \mathbb{F}_2^m$ , then

$$F_g^{-1} Z_h F_g = Z_{[h_1 A_g + h_2 C_g, h_1 B_g + h_2 D_g]} = Z_{h F_g}, \quad (21)$$

where

$$F_g = \begin{bmatrix} A_g & B_g \\ C_g & D_g \end{bmatrix}, \quad \text{and} \quad F_g^{-1} = \begin{bmatrix} D_g^\top & B_g^\top \\ C_g^\top & A_g^\top \end{bmatrix}.$$

Hence, the group generated by the symplectic transvections is normal in  $\text{Sp}(2m, \mathbb{F}_2)$ . Since  $\text{Sp}(2m, \mathbb{F}_2)$  is simple, it is generated by the set of symplectic transvections.

We shall view the row vectors that define a symplectic transvection as elements of the finite field  $\mathbb{F}_{2^m}$ . Given  $a, b, h_1, h_2 \in \mathbb{F}_{2^m}$ , let  $h = [[h_1], [h_2]]$ , then

$$\begin{aligned} Z_h([[a], [b]]) &= [[a], [b]] + ([a] \cdot [h_2] + [h_1] \cdot [b])[ [h_1], [h_2] ] \\ &\equiv (a \ b) + \text{Tr}(ah_2 + bh_1) (h_1 \ h_2) \end{aligned} \quad (22)$$

We will often write the transvection  $Z_h$  as  $Z_{(h_1, h_2)}$ , where  $h_1, h_2 \in \mathbb{F}_{2^m}$ .

### 3 The Pauli Geometry

In this section, we will define a graph on Pauli matrices, where Clifford elements act as graph automorphisms. We will build on prior work [11] showing that a set of automorphisms forms a unitary 2-design if it acts transitively on vertices, or a unitary 3-design if it acts transitively on vertices, on edges, and on non-edges.

#### 3.1 The Maximal Commutative Subgroups

To begin with, we shall review some concepts that are closely related to the Pauli group: stabilizers, maximal commutative subgroups, and stabilizer states.

A *stabilizer* group is a subgroup  $S$  of  $P_N$  generated by commuting Hermitian matrices of the form  $\pm E(a, b)$ , with the additional property that if  $E(a, b) \in S$  then  $-E(a, b) \notin S$  [21, Chapter 10]. The operators  $\frac{I_N \pm E(a, b)}{2}$  project onto the  $\pm 1$  eigenspaces of  $E(a, b)$ , respectively.

Since all elements of  $S$  are unitary, Hermitian and commute with each other, they can be diagonalized simultaneously with respect to a common orthonormal basis, and their eigenvalues are  $\pm 1$  with algebraic multiplicity  $N/2$ . We refer to such a basis as the *common eigenbasis* or simply the *eigenbasis* of the subgroup  $S$ , and to the subspace of eigenvectors with eigenvalue  $+1$  as the *+1 eigenspace* of  $S$ .

If the subgroup  $S$  is generated by  $E(a_i, b_i), i = 1, \dots, k$ , then the operator

$$\frac{1}{2^k} \prod_{i=1}^k (I_N + E(a_i, b_i)) \quad (23)$$

projects onto the  $2^{m-k}$ -dimensional subspace  $V(S)$  fixed pointwise by  $S$ , i.e., the  $+1$  eigenspace of  $S$ . The subspace  $V(S)$  is the *stabilizer code* determined by  $S$ . One uses the notation  $[[m, m-k]]$  code to represent that  $V(S)$  encodes  $m-k$  *logical* qubits into  $m$  *physical* qubits.

Let  $\psi(S)$  denote the subspace of  $\mathbb{F}_2^{2m}$  formed by the binary representations of the elements of  $S$  using the homomorphism  $\psi$  in (11). A generator matrix for  $\psi(S)$  is

$$G_S := [[a_i], [b_i]]_{i=1, \dots, k} \in \mathbb{F}_2^{k \times 2m} \text{ s.t. } G_S \Omega G_S^T = 0, \quad (24)$$

where  $0$  is the  $k \times k$  matrix with all entries zero.

A stabilizer group  $S$  defined by  $k = m$  generators is called a *maximal commutative subgroup* of  $P_N$  and  $\psi(S)$  is called a *maximal isotropic subspace* of  $\mathbb{F}_2^{2m}$ . The generator matrix  $G_S$  has rank  $m$  and can be row-reduced to  $[0 \mid I_m]$  if  $S = Z_N := \{E(0, b) : b \in \mathbb{F}_{2^m}\}$ , or to the form  $[I_m \mid P]$  if  $S$  is disjoint from  $Z_N$ . The condition  $G_S \Omega G_S^T = 0$  implies  $P = P^T$ .

*Remark 1* We will denote these maximal commutative subgroups as  $E([0 \mid I_m])$  and  $E([I_m \mid P])$ , respectively, and  $E([I_m \mid 0]) = X_N := \{E(a, 0) : a \in \mathbb{F}_{2^m}\}$ . Notice that we still employ the form of  $E(a, b)$ , where  $a, b \in \mathbb{F}_{2^m}$ , to represent elements inside of each subgroup.

### 3.2 The Kerdock Set

Kerdock sets connect Kerdock codes [25] with maximal commutative subgroups of the Pauli group. We refer the reader to [10] for more information about connections between binary and quaternary Kerdock codes and eigenbases of maximal commutative subgroups  $E([I_m | P_z])$ , where  $z \in \mathbb{F}_{2^m}$ .

We write multiplication by  $z \in \mathbb{F}_{2^m}$  as a linear transformation  $xz \equiv [x]A_z$ . For  $z = 0$ ,  $A_0 = 0$ , and for  $z = \alpha^i$  the matrix  $A_z = A^i$  for  $i = 0, 1, \dots, 2^m - 2$ , where  $A$  is the matrix that represents multiplication by the primitive element  $\alpha$ . The matrix  $A$  is the *companion matrix* of the primitive irreducible polynomial  $p(x) = p_0 + p_1x + \dots + p_{m-1}x^{m-1} + x^m$  over the binary field. Thus,

$$A := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ p_0 & p_1 & p_2 & \cdots & p_{m-1} \end{bmatrix}. \quad (25)$$

**Lemma 1** *It follows directly from the arithmetic of  $\mathbb{F}_{2^m}$  that the matrices  $A_z$  and  $W$  satisfy:*

- (a)  $A_z A_x = A_x A_z = A_{xz}$ ;
- (b)  $A_x + A_z = A_{x+z}$ ;
- (c)  $A_z W = W A_z^T$ ;

*Proof* Specifically, for (c), observe that

$$([x]A_z)W[y]^T = \text{Tr}((xz)y) = \text{Tr}(x(yz)) = [x]W([y]A_z)^T.$$

The other two properties are easily verified.  $\square$

The *Kerdock set*  $P_K(m)$  consists of all symmetric matrices  $P_z = A_z W$ , where  $z \in \mathbb{F}_{2^m}$  [10]. It is closed under binary addition, and if  $P, Q \in P_K(m)$  are distinct, then  $P + Q$  is non-singular, which in turn implies that the maximal commutative subgroups  $E([I_m | P_z])$  determined by the Kerdock matrices  $P_z$  intersect trivially. Together with  $Z_N = E([0 | I_m])$ , they partition all  $(N^2 - 1)$  non-identity Hermitian Pauli matrices. Hence, given a non-identity Hermitian Pauli matrix  $E(a, b)$ , it follows that there is a sign  $\epsilon \in \{\pm 1\}$  such that  $\epsilon E(a, b)$  is in one of the  $N + 1$  subgroups determined by all  $P_z \in P_K(m)$  and  $Z_N$ . If  $E(a, b)$  is in the maximal commutative subgroup  $E([I_m | P_z])$ , then we must have

$$[b] = [a]A_z W = [az]W = [az] \Rightarrow b = az.$$

Therefore, each element in  $E([I_m | P_z])$  or  $E([0 | I_m])$  can be written as  $E(a, az)$  or  $E(0, a)$ , respectively, for some  $a \in \mathbb{F}_{2^m}$ .

*Remark 2* Throughout this paper, we will illustrate theory developed for general  $m$  by reducing to the special case  $m = 3$ . We constructed  $\mathbb{F}_8$  by adjoining a root  $\alpha$  of  $p(x) = x^3 + x + 1$  over  $\mathbb{F}_2$ . Then, we have  $\mathbb{F}_8 = \{0, 1, \alpha, \alpha^2, \dots, \alpha^6\}$  and  $\alpha^7 = 1$ . The trace of  $0, \alpha, \alpha^2, \alpha^4$  is 0 and the trace of  $1, \alpha^3, \alpha^5, \alpha^6$  is 1.

*Example 1* It follows from (25) and (6) that

$$A_\alpha = A^1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \text{ and } W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (26)$$

We may form  $P_z = A_z W$  for any  $z \in \mathbb{F}_{2^m}$  to generate the entire Kerdock set  $P_K(m)$ .

### 3.3 The Pauli Graph

**Definition 1** The *Pauli graph*  $\mathbb{P}_N$  has  $N^2 - 1$  vertices, each of which is labeled by pairs  $\pm E(a, b)$  with  $(a, b) \neq (0, 0)$  and represented as  $(a, b)$ . A directed edge connects vertex  $(a, b)$  to vertex  $(c, d)$  if and only if  $E(a, b)$  commutes with  $E(c, d)$ , represented as a  $2 \times 2$  matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  over  $\mathbb{F}_{2^m}$ .

Apparently, if there is a directed edge connecting  $(a, b)$  to  $(c, d)$ , then  $(c, d)$  must also connect to  $(a, b)$ . The need for this deterministic additional edge between commuting Paulis will be clear when we discuss orbits in Section 4.2 and our Markov process in Section 5. Then, it follows from (8) that

$$\text{Tr} \left( \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \text{Tr}(ad + bc) = 0. \quad (27)$$

We shall distinguish two types of directed edges in  $\mathbb{P}_N$ :

1. *Type-1* edges connect vertices from the same maximal commutative subgroup  $E([I_m | P_z])$ ,  $z \in \mathbb{F}_{2^m}$ , or from  $E([0 | I_m])$ . The determinant of a type-1 edge matrix is 0:

$$\det \begin{pmatrix} a & az \\ b & bz \end{pmatrix} = 0, \text{ where } a \neq b, ab \neq 0. \quad (28)$$

2. *Type-2* edges connect vertices from different maximal commutative subgroups. The determinant of a type-2 edge matrix is a non-zero field element with trace 0:

$$\det \begin{pmatrix} a & az_1 \\ b & bz_2 \end{pmatrix} = ab(z_1 + z_2) \neq 0, \text{ where } z_1 \neq z_2, ab \neq 0. \quad (29)$$

The determinant of a  $2 \times 2$  matrix representing an ordered pair of non-commuting Pauli matrices is a non-zero field element that has trace equal to 1. We shall refer to an arbitrary  $2 \times 2$  matrix over  $\mathbb{F}_{2^m}$ , which represents an edge or a non-edge as each row corresponds to a Pauli element, as a Pauli pair matrix.

**Lemma 2** *The Pauli graph  $\mathbb{P}_N$  is a strongly regular graph with parameters*

$$n = N^2 - 1, \quad t = \frac{N^2}{2} - 2, \quad \lambda = \frac{N^2}{4} - 3, \quad \mu = \frac{N^2}{4} - 1, \quad (30)$$

where  $n$  is the number of vertices,  $t$  is the in-degree and out-degree of each vertex, and  $\lambda$  or  $\mu$  is the number of vertices joined to a pair of distinct vertices  $x, y$  according as  $x, y$  are joined or not joined respectively [26, Definiton 2.4].

*Proof* A vertex  $(c, d)$  joined to a given vertex  $(a, b)$  must satisfy  $\text{Tr}(ad + bc) = 0$  and half of the elements in  $\mathbb{F}_{2^m}$  have trace 0. For each possible value of  $ad + bc$ , there exists  $N$  pairs  $(c, d)$  that are feasible. After eliminating the solutions  $(0, 0)$  and  $(a, b)$ , we are left with  $t = \frac{N^2}{2} - 2$  distinct vertices  $(c, d)$  joined to  $(a, b)$ .

Given vertices  $(a, b)$  and  $(c, d)$ , a vertex  $(e, f)$  joined to both  $(a, b)$  and  $(c, d)$  satisfies  $\text{Tr}(af + be) = 0$  and  $\text{Tr}(cf + de) = 0$ . Half of the elements in  $\mathbb{F}_{2^m}$  have trace 0 and each pair of  $af + be$  and  $cf + de$  values have one solution for  $(e, f)$ . When  $(a, b)$  is not joined to  $(c, d)$ , we only need to eliminate the solution  $(0, 0)$ . When  $(a, b)$  is joined to  $(c, d)$  we need to eliminate  $(0, 0)$ ,  $(a, b)$  and  $(c, d)$ .  $\square$

The number of edges in  $\mathbb{P}_N$  is  $(N^2 - 1)(N^2 - 4)/2$ . The number of type-1 edges is  $(N^2 - 1)(N - 2)$  and the number of type-2 edges is  $N(N^2 - 1)(N - 2)/2$ .

Using these properties of the Pauli graph, we can now discuss unitary designs. We denote the linear operators acting on a complex Euclidean space  $\mathcal{X}$  (e.g.,  $\mathcal{X} = \mathbb{C}^N$ ) as  $L(\mathcal{X})$  and the quantum channels acting on  $L(\mathcal{X})$  as  $\mathcal{C}(\mathcal{X})$ .



**Definition 2** Let  $k$  be a positive integer and  $\mathcal{E} = \{\alpha_i, U_i\}_{i=1}^n$  be an ensemble where the unitary matrix  $U_i$  is selected with probability  $\alpha_i$ . The superoperators  $\mathcal{G}_{\mathcal{E}}, \mathcal{G}_H \in C((\mathbb{C}^N)^{\otimes k})$  are given by

$$\mathcal{G}_{\mathcal{E}}(\rho) = \sum_i \alpha_i U_i^{\otimes k} \rho (U_i^\dagger)^{\otimes k}, \quad (31)$$

$$\mathcal{G}_H(\rho) = \int_{\mathbb{U}_N} d\eta(U) U^{\otimes k} \rho (U^\dagger)^{\otimes k}, \quad (32)$$

where  $\eta(\cdot)$  represents the Haar measure on the unitary group  $\mathbb{U}_N$ . The ensemble  $\mathcal{E}$  is a *unitary  $k$ -design* if and only if  $\mathcal{G}_{\mathcal{E}} = \mathcal{G}_H$ . The linear transformations determined by  $\mathcal{G}_{\mathcal{E}}$  and  $\mathcal{G}_H$  are called  *$k$ -fold twirls*. A unitary  $k$ -design is defined by the property that the ensemble twirl coincides with the full unitary twirl.

Notice that elements of the Clifford group act by conjugation on  $P_N$ , inducing automorphisms of the graph  $\mathbb{P}_N$ . It is well known that the symplectic group  $\text{Sp}(2m, \mathbb{F}_2)$  acts transitively on vertices, on edges, and on non-edges of  $\mathbb{P}_N$ .

Following Webb [11], we say that an ensemble  $\mathcal{E} = \{\alpha_i, U_i\}_{i=1}^n$  of Clifford elements  $U_i$  is *Pauli mixing* if for every vertex  $(a, b)$  the distribution  $\{\alpha_i, U_i E(a, b) U_i^\dagger\}$  is uniform over vertices of  $\mathbb{P}_N$ . The ensemble  $\mathcal{E}$  is *Pauli 2-mixing* if it is Pauli mixing and if for every edge (resp. non-edge)  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , the distribution

$$\{\alpha_i, (U_i E(a, b) U_i^\dagger, U_i E(c, d) U_i^\dagger)\}$$

is uniform over edges (resp. non-edges) of  $\mathbb{P}_N$ .

**Theorem 1** *Let  $G$  be a subgroup of the Clifford group containing all  $D(a, b) \in HW_N$ , and let  $\mathcal{E} = \{\frac{1}{|G|}, U\}_{U \in G}$  be the ensemble defined by the uniform distribution. If  $G$  acts transitively on vertices of  $\mathbb{P}_N$ , then  $\mathcal{E}$  is a unitary 2-design, and if  $G$  acts transitively on vertices, edges and non-edges, then  $\mathcal{E}$  is a unitary 3-design.*

*Proof* Transitivity means a single orbit so that random sampling from  $G$  results in the uniform distribution on vertices, edges, and non-edges. Hence, transitivity on vertices implies  $\mathcal{E}$  is Pauli mixing and transitivity on vertices, edges and non-edges implies  $\mathcal{E}$  is Pauli 2-mixing. It now follows from [11] or [9] that Pauli mixing (resp. Pauli 2-mixing) implies  $\mathcal{E}$  is a unitary 2-design (resp. unitary 3-design).  $\square$

Theorem 1 implies that random sampling from the Clifford group gives an exact unitary 3-design.

## 4 The Unitary 2-Design

The unitary 2-design we consider in this paper is a symplectic subgroup isomorphic to  $\text{PSL}(2, 2^m)$ , and different descriptions may be found in [27], [9], and [10]. We will follow the perspective provided in [10]. We first show that it acts transitively on vertices of the graph  $\mathbb{P}_N$ . Then, we use our finite field representation to describe how this subgroup acts on ordered pairs of commuting and anti-commuting Pauli Matrices.

### 4.1 The Symplectic Subgroup Isomorphic to $\text{PSL}(2, 2^m)$

We first introduce  $\text{PSL}(2, 2^m)$  and realize each transformation as a symplectic matrix. Then, we explain why this symplectic subgroup forms a unitary 2-design by showing how the group elements permute the maximal commutative subgroups  $E([I_m | P_z])$  and  $E([0 | I_m])$  and elements within each subgroup.

The *projective special linear group* of  $2 \times 2$  matrices over  $\mathbb{F}_{2^m}$  is defined as

$$\text{PSL}(2, 2^m) := \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} : \alpha, \beta, \gamma, \delta \in \mathbb{F}_{2^m}; \alpha\delta + \beta\gamma = 1 \right\}. \quad (33)$$

The order  $|\text{PSL}(2, 2^m)| = (N + 1)N(N - 1) = 2^{3m} - 2^m$ . The action of each  $2 \times 2$  matrix  $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  over  $\mathbb{F}_{2^m}$  on 1-dimensional subspaces of  $\mathbb{F}_{2^m} \times \mathbb{F}_{2^m}$  is associated with a transformation

$$f(z) = \frac{\beta + \delta z}{\alpha + \gamma z}$$

acting on the projective line  $\mathbb{F}_{2^m} \cup \{\infty\}$ , given that

$$(1 \ z) \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = (\alpha + \gamma z \ \beta + \delta z) \equiv (0 \ 1) \text{ or } \left(1 \ \frac{\beta + \delta z}{\alpha + \gamma z}\right). \quad (34)$$

The group  $\text{PSL}(2, 2^m)$  is generated by the transformations  $z \mapsto z+x$ ,  $z \mapsto zx$ , and  $z \mapsto 1/z$ . We realize each of these transformations as a symplectic transformation. We recall that  $A_z W A_z^\top = A_z^2 W$  from part (c) of Lemma 1, and for convenience we work with maximal commutative subgroups  $E([I_m \mid A_z^2 W])$ , i.e., the Kerdock matrices are  $P_z = A_z^2 W$ . Note that every field element  $\beta \in \mathbb{F}_{2^m}$  is a square, so this is equivalent to  $P_z = A_z W$ .

(a)  $z \mapsto z+x$  becomes  $[I_m \mid A_z^2 W] \mapsto [I_m \mid A_{x+z}^2 W]$ :

$$\begin{aligned} [I_m \mid A_z^2 W] \begin{bmatrix} I_m & A_x^2 W \\ 0 & I_m \end{bmatrix} &= [I_m \mid (A_z^2 + A_x^2)W] \\ &\equiv [I_m \mid (A_{x+z}^2)W]. \end{aligned} \quad (35)$$

*Example 2* When  $m = 3$ ,  $x = \alpha$ , and  $z = \alpha^3$ , we have  $x+z = 1$ ,  $A_x = A$ , and  $A_z = A^3$ . Then,

$$A_x^2 + A_z^2 = A_{x+z}^2 = I_3 = A^0. \quad (36)$$

(b)  $z \mapsto xz$  becomes  $[I_m \mid A_z^2 W] \mapsto [I_m \mid A_{xz}^2 W]$ :

$$\begin{aligned} [I_m \mid A_z^2 W] \begin{bmatrix} A_x^{-1} & 0 \\ 0 & A_x^\top \end{bmatrix} &= [A_x^{-1} \mid A_z^2 W A_x^\top] \\ &= [A_x^{-1} \mid A_x A_z^2 W] \\ &\equiv [I_m \mid A_{xz}^2 W]. \end{aligned} \quad (37)$$

(c)  $z \mapsto 1/z$  becomes  $[I_m \mid A_z^2 W] \mapsto [I_m \mid A_{z^{-1}}^2 W]$ :

$$\begin{aligned} [I_m \mid A_z^2 W] \begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix} \begin{bmatrix} W^{-1} & 0 \\ 0 & W^\top \end{bmatrix} &= [A_z^2 W \mid I_m] \begin{bmatrix} W^{-1} & 0 \\ 0 & W \end{bmatrix} \\ &= [A_z^2 \mid W] \\ &\equiv [I_m \mid A_{z^{-1}}^2 W]. \end{aligned} \quad (38)$$

Note that if we start with  $z = 0$ , i.e., the subgroup  $E([I_m \mid 0])$ , then since  $W$  is invertible the final subgroup is  $E([0 \mid I_m])$ , interpreted as  $z = \infty$ .

Therefore,  $\text{PSL}(2, 2^m)$  is isomorphic to  $\mathfrak{P}_m$ , a group of symplectic matrices defined as

$$\mathfrak{P}_m := \langle T_{A_x^2 W}, L_{A_x^{-1}}, \Omega L_{W^{-1}}; x \in \mathbb{F}_{2^m} \rangle \cong \text{PSL}(2, 2^m), \quad (39)$$

and each  $\text{PSL}(2, 2^m)$  element induces a product of basic symplectic matrices in Table 1. The isomorphism  $\theta: \text{PSL}(2, 2^m) \rightarrow \mathfrak{P}_m$  can be defined as

$$\theta \left( \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \right) := T_{A_{\delta/\gamma}^2 W} \cdot L_{A_\gamma^{-2}} \cdot \Omega L_{W^{-1}} \cdot T_{A_{\alpha/\gamma}^2 W} \quad (40)$$

$$= \begin{bmatrix} A_\delta^2 & A_\beta^2 W \\ W^{-1} A_\gamma^2 & (A_\alpha^2)^\top \end{bmatrix}, \quad (41)$$

where  $\alpha, \beta, \gamma, \delta \in \mathbb{F}_{2^m}$  and  $\alpha\delta + \beta\gamma = 1$  [10, Lemma 23 and Corollary 24]. The induced action on maximal commutative subgroups is given by

$$E([I_m \mid A_z^2 W]) \mapsto E([I_m \mid A_{\frac{\beta + \delta z}{\alpha + \gamma z}}^2 W]) \text{ or } E([0 \mid I_m]). \quad (42)$$

Notice that the corresponding Clifford subgroup is larger than  $\text{PSL}(2, 2^m)$  since  $P_N$  forms the kernel of the homomorphism from  $\text{Cliff}_N$  to  $\text{Sp}(2m, \mathbb{F}_2)$ .

The first two factors in (40) provide transitivity on the Hermitian matrices of all maximal commutative subgroups except  $Z_N = E([0 | I_m])$ , and the last two factors enables exchanging any subgroup  $E([I_m | P_z])$  with  $E([0 | I_m])$ .

To prove that  $\text{PSL}(2, 2^m)$  acts transitively on vertices of the Pauli graph  $\mathbb{P}_N$ , we only need to show that the group is transitive on a particular subgroup, say  $E([I_m | 0])$ . For any  $(a, 0)$  and  $(b, 0)$  where  $a \neq b \in \mathbb{F}_{2^m}$ , there always exists a group element  $\begin{pmatrix} a^{-1}b & 0 \\ 0 & b^{-1}a \end{pmatrix}$  that maps  $(a, 0)$  to  $(b, 0)$ .

It then follows from Theorem 1 and (16) that random sampling from the symplectic subgroup  $\mathfrak{P}_m$  isomorphic to  $\text{PSL}(2, 2^m)$  followed by a random Pauli matrix  $D(a, b)$  produces a unitary 2-design.

However,  $\text{PSL}(2, 2^m)$  is only able to permute maximal commutative subgroups or elements within each subgroup. It is not transitive on edges of  $\mathbb{P}_N$  since it fails to mix type-1 edges and type-2 edges. Thus,  $\mathfrak{P}_m$  cannot be a unitary 3-design.

## 4.2 Orbit Invariants

$\text{PSL}(2, 2^m)$  partitions the edges of the Pauli graph into orbits, and we now identify orbit invariants.

**Definition 3** We calculate the *orbit invariant* from any representative Pauli pair matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  as follows:

- (a) The determinant  $ad + bc$  is the orbit invariant for any non-edge matrix or type-2 edge matrix.
- (b) For any type-1 edge matrix, the first row is a scalar multiple of the second row, given that the two vertices  $(a, b)$  and  $(c, d)$  are in the same maximal commutative subgroup. Its orbit invariant is the scalar  $\frac{a}{c} = \frac{b}{d} \in \mathbb{F}_{2^m} \setminus \{0, 1\}$ .

*Example 3* When  $m = 3$ , the following three Pauli pair matrices

$$E_1 = \begin{pmatrix} \alpha & 0 \\ 1 & \alpha^2 \end{pmatrix}, E_2 = \begin{pmatrix} \alpha & 0 \\ 1 & \alpha \end{pmatrix}, \text{ and } E_3 = \begin{pmatrix} \alpha & 0 \\ 1 & 0 \end{pmatrix} \quad (43)$$

have orbit invariants  $\alpha^3$ ,  $\alpha^2$ , and  $\alpha$ . According to Definition 3, (27), (28), and (29), we can recognize that  $E_1$  is a non-edge,  $E_2$  is a type-2 edge, and  $E_3$  is a type-1 edge.

**Theorem 2** Consider  $\text{PSL}(2, 2^m)$  acting on Pauli pair matrices by right multiplication. Two Pauli pair matrices are in the same orbit if and only if they have the same orbit invariant.

*Proof* Every matrix in  $\text{PSL}(2, 2^m)$  has determinant 1, so Pauli pair matrices in the same orbit share the same determinant. If the determinant is 0, then one row is a scalar multiple of the other and the scalar relation between rows is preserved by any linear transformation.

Consider two matrices, either with the same non-zero determinant, or with determinant 0 and the same scalar relation between the two rows. There always exists a linear transformation with determinant 1 that maps one to the other.  $\square$

It follows directly from Theorem 2 that we can use orbit invariants to represent and differentiate the orbits. We give some statistics about these orbits below:

- (a) There are  $\frac{N}{2}$  non-edge orbits and there are an equal number of finite field elements with trace 1. Each orbit has  $(N^2 - 1)N$  elements.
- (b) There are  $\frac{N-2}{2}$  type-2 edge orbits and there are an equal number of non-zero finite field elements with trace 0. Each orbit also has  $(N^2 - 1)N$  elements.
- (c) There are  $N - 2$  type-1 edge orbits and there are an equal number of field elements in  $\mathbb{F}_{2^m} \setminus \{0, 1\}$ . Each orbit has  $N^2 - 1$  elements.

*Example 4* For  $m = 3$ , there are 4 non-edge orbits whose invariants are  $1, \alpha^3, \alpha^5$ , and  $\alpha^6$ ; 3 type-2 edge orbits whose invariants are  $\alpha, \alpha^2$ , and  $\alpha^4$ ; and 6 type-1 edge orbits whose invariants are  $\alpha, \alpha^2, \dots, \alpha^6$ .

## 5 The Transvection Markov Process

We define a Markov process by applying a sequence of transvections to mix orbits, and a final  $\text{PSL}(2, 2^m)$  element to mix edges or non-edges within each orbit. We claim that it gives an approximate unitary 3-design by showing convergence to the uniform distribution on edges and on non-edges, in addition to the transitivity on vertices (Section 4.1).

Let  $K = (N^2 - 1)(N^2 - 4)/2$  be the number of edges and  $K' = (N^2 - 1)N^2/2$  be the number of non-edges in  $\mathbb{P}_N$ . Consider the underlying Markov chain on directed edges (resp. non-edges) with a  $K \times K$  (resp.  $K' \times K'$ ) transition matrix. Since transvections generate the full Clifford group and Clifford elements act transitively on edges and non-edges, the uniform distribution on all edges (resp. non-edges) is stationary. We are interested in the rates at which the two Markov processes converge to their corresponding stationary distributions.

Sampling  $\text{PSL}(2, 2^m)$  elements results in uniform probabilities within orbits. Therefore, it suffices to reduce the two underlying Markov chains to orbits and only consider how non-identity (symplectic) transvections transfer probability mass within the reduced state space. The dimensions of the new transition matrix on edge (resp. non-edge) orbits are  $\frac{3}{2}(N - 2) \times \frac{3}{2}(N - 2)$  (resp.  $\frac{N}{2} \times \frac{N}{2}$ ).

### 5.1 The Transvection Markov Chain on Non-Edges

**Theorem 3** *Consider the Markov process with state space consisting of all non-edge orbits. The matrix  $Q_1$  of state transition probabilities is given by*

$$Q_1 = \frac{1}{4(N^2 - 1)} [(N^2 - 4)I_{N/2} + 6NJ_{N/2}] \in \mathbb{R}^{\frac{N}{2} \times \frac{N}{2}}, \quad (44)$$

where  $I_{N/2}, J_{N/2} \in \mathbb{R}^{\frac{N}{2} \times \frac{N}{2}}$ ,  $I_{N/2}$  is the identity matrix, and  $J_{N/2}$  is the all ones matrix.

*Proof* We apply a random transvection  $Z_{(h_1, h_2)}$ , where  $h_1, h_2 \in \mathbb{F}_{2^m}$  and  $(h_1, h_2) \neq (0, 0)$ , to a non-edge matrix  $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$  with orbit invariant  $ab$ , where  $\text{Tr}(ab) = 1$ . According to (22), we have the following four cases:

- (a) Applying a transvection with  $\text{Tr}(ah_2) = 0$  and  $\text{Tr}(bh_1) = 0$  fixes the non-edge:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}. \quad (45)$$

There are two constraints on  $h_1$  and  $h_2$ , also  $(h_1, h_2) \neq (0, 0)$ . Therefore, the number of possible transvections is  $\left(\frac{N}{2}\right)^2 - 1 = \frac{N^2}{4} - 1$ .

- (b) Applying a transvection with  $\text{Tr}(ah_2) = 1$  and  $\text{Tr}(bh_1) = 1$ , we obtain

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a + h_1 & h_2 \\ h_1 & b + h_2 \end{pmatrix} \quad (46)$$

and the new orbit invariant is

$$(a + h_1)(b + h_2) + h_1h_2 = ab + ah_2 + bh_1.$$

Since  $\text{Tr}(ab) + \text{Tr}(ah_2) + \text{Tr}(bh_1) = 1$ , the resulting Pauli pair is not an edge. The products  $ah_2$  and  $bh_1$  range over all field elements with trace 1. Given a field element  $x$  with  $\text{Tr}(x) = 1$ , the number of solutions to  $x = ab + ah_2 + bh_1$  is simply the number of solutions to  $ah_2 + bh_1 = 0$ . There are  $N/2$  transvections to each of the  $N/2$  orbits.

- (c) Applying a transvection with  $\text{Tr}(ah_2) = 0$  and  $\text{Tr}(bh_1) = 1$ , we obtain

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a & 0 \\ h_1 & b + h_2 \end{pmatrix} \quad (47)$$

and the new orbit invariant is

$$a(b + h_2) = ab + ah_2.$$

Since  $\text{Tr}(ab) + \text{Tr}(ah_2) = 1$ , the resulting Pauli pair is not an edge. A similar argument to that used in part (b) shows that there are  $N/2$  transvections to each of the  $N/2$  orbits.

(d) Applying a tranvection with  $\text{Tr}(ah_2) = 1$  and  $\text{Tr}(bh_1) = 0$ , we obtain

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a + h_1 & h_2 \\ 0 & b \end{pmatrix} \quad (48)$$

and the new orbit invariant is

$$(ah_1)b = ab + bh_1.$$

The same argument used in part (c) shows that there are  $N/2$  transvections to each of the  $N/2$  orbits.

There are  $N^2 - 1$  transvections, and each is a symplectic matrix that preserves non-edges in the Pauli graph. Case (a) contributes to the diagonal component  $\frac{N^2-4}{4(N^2-1)}I_{N/2}$  in  $Q_1$  and cases (b), (c), and (d) contribute to  $\frac{6N}{4(N^2-1)}J_{N/2}$  in  $Q_1$ .  $\square$

## 5.2 The Transvection Markov Chain on Edges

**Theorem 4** Consider the Markov process with state space consisting of all edge orbits. Set  $M_1 = N - 2$  and  $M_2 = \frac{N-2}{2}$ . Index the first  $M_1$  rows and columns of the state transition matrix  $Q_0$  by the type-1 orbits, and the remaining  $M_2$  rows and columns by the type-2 orbits. Then,  $Q_0$  is given by

$$Q_0 = \frac{1}{4(N^2 - 1)} \begin{bmatrix} (N^2 - 4)I_{M_1} & NR^T \\ R & (N^2 - 4)I_{M_2} + 6NJ_{M_2} \end{bmatrix} \quad (49)$$

where  $I_{M_1}$  and  $I_{M_2}$  are the identity matrices,  $J_{M_2}$  is the all ones matrix,  $R$  is non-negative, each row sum of  $R$  is  $6N$ , and each column sum of  $R$  is  $3N$ .

*Proof* We determine the lower right block of the transition matrix  $Q_0$  by making a slight modification to the proof of Theorem 3. Starting with  $\text{Tr}(ab) = 0$ , we consider  $x$  with  $\text{Tr}(x) = 0$  in cases (b), (c), and (d). In each case, if  $x = 0$ , we get a type-1 orbit and there are  $N/2$  such transvections. Thus, the matrix  $R/4$  in the lower left block has row sum  $3N/2$ . If  $x \neq 0$ , of which there are  $(N - 2)/2$  cases, we get a transition to one of the  $(N - 2)/2$  type-2 orbits and there are still  $N/2$  such transvections in each case. This contributes to the  $J_{M_2}$  term in the lower right block. Finally, case (a) from Theorem 3 produces the identity component.

We now start with a type-1 edge  $\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}$ , where  $a \neq b$ ,  $ab \neq 0$ , and has orbit invariant  $\frac{a}{b}$ , and apply a random transvection  $Z_{(h_1, h_2)}$ , where  $h_1, h_2 \in \mathbb{F}_2^m$  and  $(h_1, h_2) \neq (0, 0)$ . By distinguishing four cases similar to the proof of Theorem 3, we notice that the upper left block of  $Q_0$ , which describes the probability of transitioning from type-1 orbits to type-1 orbits by transvections, contains only the diagonal component  $\frac{N^2-4}{4(N^2-1)}I_{M_1}$ .

Since transvections are self-inverse, the upper right block of  $Q_0$  must be some scalar multiple of  $R^T$ . The transvection Markov chain on edges is irreducible, so by the Perron-Frobenius Theorem, there is a unique stationary distribution. Since the uniform distribution on edges is stationary, we observe that

$$\mathbf{w}_1 = \left[ \frac{1}{N} \dots \frac{1}{N} 1 \dots 1 \right] \in \mathbb{R}^{N-2+\frac{N-2}{2}} \quad (50)$$

is the stationary distribution of the Markov chain on edge orbits. Given that the row sum of  $R$  is  $6N$ , the upper right block is  $NR^T$ .  $\square$

*Example 5* Here  $m = 3$ , and we derive the matrix  $R$ . Starting with a type-2 edge  $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$  for which the orbit invariant is  $ab = \alpha$ , we consider the following three cases:

(a) Applying a transvection with  $\text{Tr}(ah_2) = 1$  and  $\text{Tr}(bh_1) = 1$ , we obtain

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a + h_1 & h_2 \\ h_1 & b + h_2 \end{pmatrix}, \quad (51)$$

with determinant

$$(a + h_1)(b + h_2) + h_1h_2 = ab + ah_1 + bh_2.$$

When  $ah_1 + bh_2 = ab = \alpha$ , the resulting Pauli pair is a type-1 edge with orbit invariant

$$\frac{a + h_1}{h_1} = \frac{ab + bh_1}{bh_1} = \frac{\alpha}{bh_1} + 1.$$

Since  $bh_1$  takes the values  $1, \alpha^3, \alpha^5$ , or  $\alpha^6$  with equal probability, the corresponding orbit invariants take the values

$$\frac{\alpha}{1} + 1 = \alpha^3, \quad \frac{\alpha}{\alpha^3} + 1 = \alpha^4, \quad \frac{\alpha}{\alpha^5} + 1 = \alpha, \quad \text{or} \quad \frac{\alpha}{\alpha^6} + 1 = \alpha^6 \quad (52)$$

with equal probability.

(b) Applying a transvection with  $\text{Tr}(ah_2) = 0$  and  $\text{Tr}(bh_1) = 1$ , we obtain

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{Z_{(h_1, h_2)}} \begin{pmatrix} a & 0 \\ h_1 & b + h_2 \end{pmatrix} \quad (53)$$

with determinant

$$a(b + h_2) = ab + ah_2.$$

When  $ah_2 = ab = \alpha$ , the resulting Pauli pair is a type-1 edge with orbit invariant

$$\frac{a}{h_1} = \frac{ab}{bh_1} = \frac{\alpha}{bh_1}.$$

Again,  $bh_1$  takes the values  $1, \alpha^3, \alpha^5$ , or  $\alpha^6$  with equal probability, so the corresponding orbit invariants take the values

$$\frac{\alpha}{1} = \alpha, \quad \frac{\alpha}{\alpha^3} = \alpha^5, \quad \frac{\alpha}{\alpha^5} = \alpha^3, \quad \text{or} \quad \frac{\alpha}{\alpha^6} = \alpha^2 \quad (54)$$

with equal probability.

(c) By symmetry, the result of applying a transvection with  $\text{Tr}(ah_2) = 1$  and  $\text{Tr}(bh_1) = 0$  is the same as part (b).

A type-2 edge with orbit invariant  $\alpha$  is equally likely to transition to any of the type-1 orbits. The same conclusion holds for a type-2 edge with orbit invariant  $\alpha^2$  or  $\alpha^4$ , and so

$$R_3 = \begin{matrix} & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \\ \alpha & \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \\ \alpha^2 & \\ \alpha^4 & \end{matrix} \times 8. \quad (55)$$

*Remark 3* When  $m = 3$ , the matrix  $R$  is a scalar multiple of a all ones matrix, but for general  $m$ , it may not be the case that a type-2 orbit is equally likely to transition to all type-1 orbits. For example, when  $m = 4$  and  $\alpha$  is a root of  $x^4 + x + 1$ . Suppose we start with a type-2 edge  $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$  for which the orbit invariant is  $ab = \alpha$ , then the number of possible transvections to achieve each type-1 orbits is

$$\alpha \begin{matrix} \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 & \alpha^8 & \alpha^9 & \alpha^{10} & \alpha^{11} & \alpha^{12} & \alpha^{13} & \alpha^{14} \\ \begin{bmatrix} 1 & 2 & 1 & 1 & 3 & 2 & 2 & 2 & 2 & 3 & 1 & 1 & 2 & 1 \end{bmatrix} \end{matrix} \quad (56)$$

### 5.3 Eigenvectors and Eigenvalues

We first discuss the eigenvectors and eigenvalues of  $Q_1$ . The all-one vector  $\mathbf{x}_1$  of length  $N/2$  is an eigenvector of  $Q_1$  with eigenvalue 1, i.e.,  $\mathbf{x}_1 Q_1 = \mathbf{x}_1$ . Any vector orthogonal to  $\mathbf{x}_1$  is an eigenvector of  $J_{N/2}$  with eigenvalue 0. Then, for any vector  $\mathbf{x}_2$  such that  $\mathbf{x}_2 J_{N/2} = 0$ , we have

$$\mathbf{x}_2 Q_1 = \frac{N^2 - 4}{4(N^2 - 1)} \mathbf{x}_2. \quad (57)$$

Therefore, the second largest eigenvalue of  $Q_1$  is

$$\lambda_{Q_1} = \frac{N^2 - 4}{4(N^2 - 1)} \quad (58)$$

with multiplicity  $\frac{N}{2} - 1$ .

We proceed to discuss the eigenvectors and eigenvalues of  $Q_0$ . Define

$$J' = \frac{1}{4(N^2 - 1)(N - 2)} \begin{bmatrix} (N^2 - 4)J_{M_1} & 6N^2 J_{M_1 \times M_2} \\ 6N J_{M_2 \times M_1} & (8N^2 - 12N - 8)J_{M_2} \end{bmatrix}, \quad (59)$$

where  $J_{M_1 \times M_2}$  and  $J_{M_2 \times M_1}$  are the  $M_1 \times M_2$  and  $M_2 \times M_1$  all ones matrices. By direct calculation we obtain

$$J' Q_0 = Q_0 J' = J' J'. \quad (60)$$

Then, for any eigenvector  $\mathbf{w}$  of  $J'$  with eigenvalue  $\lambda$ , we have

$$\mathbf{w} J' Q_0 = \mathbf{w} J' J' \Rightarrow \lambda \mathbf{w} Q_0 = \lambda^2 \mathbf{w} \Rightarrow \mathbf{w} Q_0 = \lambda \mathbf{w}. \quad (61)$$

Therefore,  $\mathbf{w}_1$  in (50) and

$$\mathbf{w}_2 = [1 \dots 1 \ -2 \dots -2] \in \mathbb{R}^{N-2+\frac{N-2}{2}} \quad (62)$$

are two (left) eigenvectors for both  $J'$  and  $Q_0$  with corresponding eigenvalues 1 and  $\frac{N^2 - 6N - 4}{4(N^2 - 1)}$ . The eigenvector  $\mathbf{w}_1$  is the stationary distribution of  $Q_0$ .

Additionally, for any eigenvector  $\mathbf{v}$  of  $Q_0$  with eigenvalue  $\lambda$ , we have

$$\mathbf{v} \mathbf{J}' \mathbf{J}' = \mathbf{v} Q_0 J' = \lambda \mathbf{v} J'. \quad (63)$$

Thus, either  $\mathbf{v} J' = 0$  or  $\mathbf{v} J'$  is an eigenvector of  $J'$  with eigenvalue  $\lambda$ . It remains to calculate the eigenvalues associated with eigenvectors  $\mathbf{v}$  of  $Q_0$  that satisfy  $\mathbf{v} J' = 0$ . The action of  $Q_0$  on these eigenvectors is given by the matrix

$$Q'_0 = \frac{1}{4(N^2 - 1)} \begin{bmatrix} (N^2 - 4)I_{M_1} & NR^\top \\ R & (N^2 - 4)I_{M_2} \end{bmatrix}. \quad (64)$$

Setting

$$R' = \frac{1}{4(N^2 - 1)} \begin{bmatrix} 0 & NR^\top \\ R & 0 \end{bmatrix}, \quad (65)$$

we have

$$R'^2 = \frac{1}{16(N^2 - 1)^2} \begin{bmatrix} NR^\top R & 0 \\ 0 & NRR^\top \end{bmatrix}. \quad (66)$$

Since the row sum of  $R$  is  $6N$  and the column sum is  $3N$ , the largest singular value of  $R$  is  $3\sqrt{2N}$ . Therefore, the largest eigenvalue of  $R'^2$  is  $\frac{N(3\sqrt{2N})^2}{16(N^2 - 1)^2}$ . Then, the largest eigenvalue of  $R'$  is  $\frac{3\sqrt{2N}\sqrt{N}}{4(N^2 - 1)}$ . If  $\lambda_{Q_0}$  is the second largest eigenvalue of  $Q_0$ , then

$$\lambda_{Q_0} < \frac{N^2 - 4 + 3N\sqrt{2N}}{4(N^2 - 1)} \quad (67)$$

and the minimum eigenvalue  $\lambda_{\min, Q_0}$  of  $Q_0$  is positive when  $m \geq 5$  (i.e.,  $N \geq 32$ ) since

$$\lambda_{\min, Q_0} > \frac{N^2 - 4 - 3N\sqrt{2N}}{4(N^2 - 1)} > 0. \quad (68)$$

## 6 Convergence Analysis

In this section, we prove that the transvection Markov process gives an  $\epsilon$ -approximate unitary 3-design and analyze its convergence rate using the second largest eigenvalues obtained in Section 5.3. The proof is also inspired by Section 6 in [28] and how Webb [11] proved that the Pauli 2-mixing forms an exact unitary 3-design.

We now describe the general idea of our proof. The diamond norm difference, which can be upper-bounded by  $l_2$  norm, tells us how “approximate” our design  $\mathcal{G}_\mathcal{E}$  is to an exact unitary 3-design  $\mathcal{G}_H$ . For an exact 3-design, we know it acts transitively on edges and non-edges. For our approximate 3-design, we have the transition matrices  $Q_1$  and  $Q_0$  that describe how it acts on edges and non-edges. Therefore, we can expand  $\mathcal{G}_\mathcal{E}$  and  $\mathcal{G}_H$  using Pauli elements and transition probabilities. Since the second largest eigenvalues of  $Q_1$  and  $Q_0$  tell us how fast the Markov processes converge, the transition probabilities can be accurately characterized.

### 6.1 The Approximate Unitary $k$ -Designs

**Definition 4** ([29, Chapter 9.1.6], [28, Definition 2.4]) For quantum channels  $\mathcal{N}, \mathcal{M} \in C(\mathcal{X})$ , the *diamond-norm distance* is given by

$$\|\mathcal{N} - \mathcal{M}\|_\diamond = \sup_\rho \|(\mathcal{N} \otimes I_R)(\rho) - (\mathcal{M} \otimes I_R)(\rho)\|_1 \quad (69)$$

where  $\dim(R) = \dim(\mathcal{X})$ . The diamond norm tells us the distinguishability of two quantum channels in an operational sense.

**Definition 5** ([28, Definition 2.5])  $\mathcal{G}_\mathcal{E}$  is an  $\epsilon$ -approximate unitary  $k$ -design if

$$\|\mathcal{G}_\mathcal{E} - \mathcal{G}_H\|_\diamond \leq \epsilon. \quad (70)$$

**Theorem 5** *Random sampling of  $O(\log(N^5/\epsilon))$  transvections, followed by a random element from  $\mathfrak{P}_m \cong PSL(2, 2^m)$  and a random Pauli matrix  $D(a, b)$  gives a  $\epsilon$ -approximate unitary 3-design.*

We will use the rest of Section 6 to prove Theorem 5. Denote  $\overline{P}_N$  as the set of all  $E(a, b)$  Pauli matrices. Since  $\overline{P}_N$  forms an orthogonal basis for  $L(\mathbb{C}^N)$ , each linear operator  $\rho \in L((\mathbb{C}^N)^{\otimes 6})$  can be written as

$$\rho = \frac{1}{N^3} \sum_{\substack{p_i \in \overline{P}_N, \\ i \in \{1, \dots, 6\}}} \Gamma(p_1, \dots, p_6) p_1 \otimes \dots \otimes p_6, \quad (71)$$

where

$$\Gamma(p_1, \dots, p_6) = \frac{1}{N^3} \text{Tr}((p_1 \otimes \dots \otimes p_6) \rho) \in \mathbb{R} \quad (72)$$

and

$$\sum_{\substack{p_i \in \overline{P}_N, \\ i \in \{1, \dots, 6\}}} \Gamma^2(p_1, \dots, p_6) \leq 1. \quad (73)$$

Because of this nice expansion into Pauli elements, we only need to understand the effect of  $\mathcal{G}_\mathcal{E}$  and  $\mathcal{G}_H$  acting on each element of  $\overline{P}_N^{\otimes 3}$ . Notice that Theorem 5 gives an exact unitary 2-design since it still acts transitively on vertices. Then, it follows from Webb [11, Proof of Lemma 4] that when  $\rho = p_1 \otimes p_2 \otimes I$  or  $\rho = p_1 \otimes p_2 \otimes p_3$  with  $p_1 p_2 p_3 \not\propto I$  (up to permutations of the underlying registers), we have

$$\mathcal{G}_\mathcal{E}(\rho) = \mathcal{G}_H(\rho). \quad (74)$$

Therefore, we only need to consider the case when  $\rho = p_1 \otimes p_2 \otimes p_1 p_2$  and  $p_1 \neq p_2$ :

$$\|\mathcal{G}_\mathcal{E} - \mathcal{G}_H\|_\diamond^2 = \sup_\rho \|(\mathcal{G}_\mathcal{E} \otimes I_R)(\rho) - (\mathcal{G}_H \otimes I_R)(\rho)\|_1^2$$



$$\begin{aligned}
&\leq N^6 \sup_{\rho} \|(\mathcal{G}_{\mathcal{E}} \otimes I_R)(\rho) - (\mathcal{G}_H \otimes I_R)(\rho)\|_2^2 \\
&= \sup_{\rho} \left\| \sum_{p_i \in \overline{\mathcal{P}}_N} \Gamma(p_1, \dots, p_6) \left[ \mathcal{G}_{\mathcal{E}}(p_1 \otimes p_2 \otimes p_3) \otimes p_4 \otimes p_5 \otimes p_6 \right. \right. \\
&\quad \left. \left. - \mathcal{G}_H(p_1 \otimes p_2 \otimes p_3) \otimes p_4 \otimes p_5 \otimes p_6 \right] \right\|_2^2 \\
&= \sup_{\rho} \left\| \sum_{\substack{p_i \in \overline{\mathcal{P}}_N \\ p_3 = p_1 p_2, p_1 \neq p_2}} \Gamma(p_1, \dots, p_6) \left[ \mathcal{G}_{\mathcal{E}}(p_1 \otimes p_2 \otimes p_1 p_2) \right. \right. \\
&\quad \left. \left. - \mathcal{G}_H(p_1 \otimes p_2 \otimes p_1 p_2) \right] \otimes p_4 \otimes p_5 \otimes p_6 \right\|_2^2. \tag{75}
\end{aligned}$$

Define

$$\mathcal{E}_{(p_1, p_2) \rightarrow (q_1, q_2)} = \{(\alpha, U) \in \mathcal{E} : U p_1 U^\dagger = q_1, U p_2 U^\dagger = q_2\}. \tag{76}$$

Suppose the transition matrices for the Markov chains of sampling random transvections on all edges and non-edges are  $P_0$  and  $P_1$  respectively. Then, if we start from the Pauli pair  $(p_1, p_2)$ , the probability of reaching  $(q_1, q_2)$  after  $t$  steps on the chain is

$$g_t^{[p_1, p_2]}(q_1, q_2; p_1, p_2) = \mathbf{e}_{(p_1, p_2)} P_{[p_1, p_2]}^t \mathbf{e}_{(q_1, q_2)}^\top, \tag{77}$$

where  $[p_1, p_2] = 0$  if  $p_1$  and  $p_2$  commute, otherwise  $[p_1, p_2] = 1$  and  $\mathbf{e}_{(p_1, p_2)}$  is a unit row vector where only position  $(p_1, p_2)$  is non-zero. Then, for the ensemble  $\mathcal{E}$  generated after  $t$  steps on the corresponding Markov chain, we have

$$\begin{aligned}
\mathcal{G}_{\mathcal{E}}(p_1 \otimes p_2 \otimes p_1 p_2) &= \sum_{(\alpha, U) \in \mathcal{E}} \alpha U p_1 U^\dagger \otimes U p_2 U^\dagger \otimes U p_1 U^\dagger U p_2 U^\dagger \\
&= \sum_{\substack{q_1, q_2 \in \overline{\mathcal{P}}_N : \\ [q_1, q_2] = [p_1, p_2]}} \sum_{(\alpha, U) \in \mathcal{E}_{(p_1, p_2) \rightarrow (q_1, q_2)}} \alpha q_1 \otimes q_2 \otimes q_1 q_2 \\
&= \sum_{\substack{q_1, q_2 \in \overline{\mathcal{P}}_N : \\ [q_1, q_2] = [p_1, p_2]}} g_t^{[p_1, p_2]}(q_1, q_2; p_1, p_2) q_1 \otimes q_2 \otimes q_1 q_2 \tag{78}
\end{aligned}$$

and

$$\mathcal{G}_H(p_1 \otimes p_2 \otimes p_1 p_2) = \sum_{\substack{q_1, q_2 \in \overline{\mathcal{P}}_N : \\ [q_1, q_2] = [p_1, p_2]}} g_\infty^{[p_1, p_2]}(q_1, q_2; p_1, p_2) q_1 \otimes q_2 \otimes q_1 q_2. \tag{79}$$

Then, by continuing (75) and following the orthogonality of the Pauli operators under the Hilbert-Schmidt inner product, we obtain

$$\begin{aligned}
\|\mathcal{G}_{\mathcal{E}} - \mathcal{G}_H\|_\diamond^2 &= \sup_{\rho} \left\| \sum_{\substack{p_i \in \overline{\mathcal{P}}_N \\ [p_1, p_2] = 0}} \Gamma(p_1, \dots, p_6) \sum_{\substack{q_1, q_2 \in \overline{\mathcal{P}}_N : \\ [q_1, q_2] = 0}} \left[ g_t^0(q_1, q_2; p_1, p_2) \right. \right. \\
&\quad \left. \left. - g_\infty^0(q_1, q_2; p_1, p_2) \right] q_1 \otimes q_2 \otimes q_1 q_2 \otimes p_4 \otimes p_5 \otimes p_6 \right. \\
&\quad \left. + \sum_{\substack{p_i \in \overline{\mathcal{P}}_N \\ [p_1, p_2] = 1}} \Gamma(p_1, \dots, p_6) \sum_{\substack{q_1, q_2 \in \overline{\mathcal{P}}_N : \\ [q_1, q_2] = 1}} \left[ g_t^1(q_1, q_2; p_1, p_2) \right. \right.
\end{aligned}$$

$$\begin{aligned}
& - g_\infty^1(q_1, q_2; p_1, p_2) \Big] q_1 \otimes q_2 \otimes q_1 q_2 \otimes p_4 \otimes p_5 \otimes p_6 \Big\|_2^2 \\
= & N^6 \sup_\rho \left[ \sum_{[p_1, p_2]=0} \Gamma^2(\dots) \sum_{[q_1, q_2]=0} (g_t^0(\dots) - g_\infty^0(\dots))^2 \right. \\
& \left. + \sum_{[p_1, p_2]=1} \Gamma^2(\dots) \sum_{[q_1, q_2]=1} (g_t^1(\dots) - g_\infty^1(\dots))^2 \right]. \tag{80}
\end{aligned}$$

We shall show in Section 6.2 that when  $t = O(\log(N^5/\epsilon))$ , using the second largest eigenvalues of  $Q_0$  and  $Q_1$  defined in (49) and (44), we have

$$\sum_{[q_1, q_2]=0} (g_t^0(\dots) - g_\infty^0(\dots))^2 < \left(\frac{\epsilon}{N^3}\right)^2 \tag{81}$$

and

$$\sum_{[q_1, q_2]=1} (g_t^1(\dots) - g_\infty^1(\dots))^2 < \left(\frac{\epsilon}{N^3}\right)^2. \tag{82}$$

Therefore, we can prove Theorem 5 by further simplifying (80) into

$$\|\mathcal{G}_\mathcal{E} - \mathcal{G}_H\|_\diamond^2 \leq N^6 \sup_\rho \sum \Gamma^2(\dots) \left(\frac{\epsilon}{N^3}\right)^2 \leq \epsilon^2. \tag{83}$$

## 6.2 Orbits Mixing Time

We first find the connection between the probability distribution of the Markov chain on edges and that of edge orbits. Suppose

$$\mathbf{x} = [x_1, \dots, x_{N-2}, x_{N-1}, \dots, x_{\frac{3}{2}(N-2)}] \tag{84}$$

is a probability distribution on all edge orbits. Sampling Clifford elements will distribute the probability mass equally among all the edges within each orbit. Since each type-1 orbit has  $N^2 - 1$  elements and each type-2 orbit has  $(N^2 - 1)N$  elements, we define the corresponding probability distribution on all edges after sampling a random Clifford element as

$$f(\mathbf{x}) = \left[ \frac{x_1}{N^2 - 1}, \dots, \frac{x_{N/2}}{(N^2 - 1)N}, \dots \right], \tag{85}$$

where each entry  $\frac{x_i}{N^2 - 1}$  is repeated  $N^2 - 1$  times, for  $1 \leq i \leq (N - 2)$ , and each entry  $\frac{x_i}{(N^2 - 1)N}$  is repeated  $(N^2 - 1)N$  times, for  $i \geq N - 1$ . Then, we see that

$$\begin{aligned}
\|f(\mathbf{x}) - f(\mathbf{y})\|_1 &= \sum_{i=1}^{(N^2-1)(N^2-4)/2} |f(\mathbf{x})_i - f(\mathbf{y})_i| \\
&= \sum_{i=1}^{N-2} (N^2 - 1) \left| \frac{x_i}{N^2 - 1} - \frac{y_i}{N^2 - 1} \right| \\
&\quad + \sum_{i=N-1}^{\frac{3}{2}(N-2)} (N^2 - 1)N \left| \frac{x_i}{(N^2 - 1)N} - \frac{y_i}{(N^2 - 1)N} \right| \\
&= \sum_{i=1}^{\frac{3}{2}(N-2)} |x_i - y_i| = \|\mathbf{x} - \mathbf{y}\|_1. \tag{86}
\end{aligned}$$

**Definition 6** Let  $Q$  be the transition matrix of an irreducible and aperiodic Markov process whose stationary distribution is  $\pi$ . The mixing time  $\tau$  is given by

$$\tau(\epsilon) = \max_{\mathbf{s}: \|\mathbf{s}\|_1=1} \min_t \{t \geq 0: \|\mathbf{s}Q^t - \pi\|_1 \leq \epsilon\}. \quad (87)$$

**Theorem 6 ([28, Theorem 4.5])** *The mixing time can be bounded above as*

$$\tau(\epsilon) \leq \frac{1}{\Delta} \ln \frac{1}{\pi_* \epsilon}, \quad (88)$$

where  $\pi_* = \min \pi(x)$  and  $\Delta = \min(1 - \lambda_2, 1 + \lambda_{\min})$ . Here,  $\lambda_2$  is the second largest eigenvalue and  $\lambda_{\min}$  is the smallest.

Therefore, following from (50) and (67), the mixing time for  $Q_0$  is bounded by

$$\begin{aligned} t_{\min} = \tau\left(\frac{\epsilon}{N^3}\right) &\leq \frac{1}{1 - \lambda_{Q_0}} \ln \frac{1}{\min\left(\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|_1}\right) \frac{\epsilon}{N^3}} \\ &\leq \frac{1}{1 - \frac{N^2-4+3N\sqrt{2N}}{4(N^2-1)}} \ln \frac{1}{\frac{\epsilon}{N^3} \left(\frac{1}{N}\right) / \left(\frac{N^2-4}{2N}\right)} \\ &\approx \frac{1}{1 - 1/4} \ln \frac{N^3(N^2-4)}{2\epsilon} = O(\log(N^5/\epsilon)) \end{aligned} \quad (89)$$

[30] and  $t_{\min}$  is larger than the mixing time of  $Q_1$  defined in (44). Since

$$\|f(\mathbf{x}) - f(\mathbf{y})\|_2 \leq \|f(\mathbf{x}) - f(\mathbf{y})\|_1 = \|\mathbf{x} - \mathbf{y}\|_1, \quad (90)$$

both (81) and (82) can be upper bounded by  $\left(\frac{\epsilon}{N^3}\right)^2$ . Theorem 5 is thus proved.

## 7 Conclusion

In this paper, we have proved that we can obtain an  $\epsilon$ -approximate unitary 3-design by random sampling  $O(\log(N^5/\epsilon))$  transvections followed by a random Clifford  $\text{PSL}(2, 2^m)$  element and a random Pauli matrix.

The key to an exact unitary 3-design is the transitivity on the Pauli element pairs of the same commutativity. The transvection Markov process that we designed exactly enables all Pauli pairs to converge to such a uniform stationary distribution. We first start from a well-established unitary 2-design, a symplectic subgroup isomorphic to  $\text{PSL}(2, 2^m)$  based on Kerdock sets. This unitary 2-design acts transitively on Pauli elements, but partitions the Pauli pairs into multiple orbits. Using a finite field representation of Paulis, and the fact that transvections generate the symplectic group, we characterize the orbits of Pauli pairs and analyze how each transvection acts on the orbits. Finally, we analyze the convergence rate to an  $\epsilon$ -approximate 3-design using the second largest eigenvalues of the transition matrices of the “edge”- and “orbit”-Markov chains.

Interestingly, when we look at the history of related research, scientists have already discovered that the Clifford group forms a unitary 2-design early in 2001 [1]. But it was not proved until 2016 that it is also a unitary 3-design and it is the minimal 3-design except for dimension 4 [31] [11]. Our work fills in the gap by constructing the 3-design using the exact building blocks provided by the 2-design.

Given that Mølmer-Sørensen gates (i.e., 2-qubit transvections [19]) are native gates in trapped-ion systems, it would be interesting to investigate if employing only these transvections in our Markov chain suffices to produce an approximate unitary 3-design. Moreover, it might be possible to extend this Markov chain approximation framework to create unitary  $k$ -designs for  $t > 3$  by mixing the Kerdock design with, perhaps, a generalization of symplectic transvections.

## Acknowledgments

We are grateful to Tanmay Singal and Min-Hsiu Hsieh for pointing out that we needed to define the Pauli graph as a directed graph in order to accurately analyze the Transvection Markov Chain. They also communicated that it is possible to improve the convergence analysis to  $O(\log(N^3/\epsilon))$  [30].

The work of N. Rengaswamy and R. Calderbank was supported in part by the National Science Foundation (NSF) under Grant no. 1908730.

## References

1. D. DiVincenzo, D. Leung, and B. Terhal, “Quantum Data Hiding,” vol. 48, no. 3, pp. 580–598, 2002.
2. A. Roy and A. J. Scott, “Unitary designs and codes,” *Designs, Codes Cryptogr.*, vol. 53, no. 1, pp. 13–31, 2009. [Online]. Available: <http://link.springer.com/10.1007/s10623-009-9290-2>
3. Y. Nakata, C. Hirche, C. Morgan, and A. Winter, “Unitary 2-designs from random X- and Z-diagonal unitaries,” *J. Math. Phys.*, vol. 58, no. 5, p. 052203, 2017. [Online]. Available: <http://aip.scitation.org/doi/10.1063/1.4983266>
4. P. Hayden, M. Horodecki, A. Winter, and J. Yard, “A decoupling approach to the quantum capacity,” *Open Sys. Inf. Dyn.*, vol. 15, no. 01, pp. 7–19, 2008. [Online]. Available: <http://arxiv.org/abs/quant-ph/0702005>
5. O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, “Decoupling with unitary approximate two-designs,” *New J. Phys.*, vol. 15, no. 5, p. 053022, 2013. [Online]. Available: <https://doi.org/10.1088/1367-2630/15/5/053022>
6. C. Dankert, R. Cleve, J. Emerson, and E. Livine, “Exact and approximate unitary 2-designs and their application to fidelity estimation,” *Phys. Rev. A*, vol. 80, no. 1, p. 012304, 2009. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.80.012304>
7. J. Emerson, R. Alicki, and K. Życzkowski, “Scalable noise estimation with random unitary operators,” *J. Opt. B Quantum Semiclassical Opt.*, vol. 7, no. 10, pp. S347–S352, 2005. [Online]. Available: <https://doi.org/10.1088/1464-4266/7/10/021>
8. E. Magesan, J. M. Gambetta, and J. Emerson, “Characterizing quantum gates via randomized benchmarking,” *Phys. Rev. A*, vol. 85, no. 4, p. 042311, 2012. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.85.042311>
9. R. Cleve, D. Leung, L. Liu, and C. Wang, “Near-linear constructions of exact unitary 2-designs,” *Quantum Inf. Comput.*, vol. 16, no. 9–10, pp. 0721–0756, 2016, [Online]. Available: <https://arxiv.org/pdf/1501.04592.pdf>
10. T. Can, N. Rengaswamy, R. Calderbank, and H. D. Pfister, “Kerdock Codes Determine Unitary 2-Designs,” vol. 66, no. 10, pp. 6104–6120, 2020. [Online]. Available: <http://arxiv.org/abs/1904.07842>
11. Z. Webb, “The Clifford group forms a unitary 3-design,” *Quantum Inf. Comput.*, vol. 16, no. 15–16, pp. 1379–1400, 2016, [Online]. Available: <https://arxiv.org/pdf/1510.02769.pdf>
12. A. Salam, E. Al-Aidarous, and A. E. Farouk, “Optimal symplectic Householder transformations for SR decomposition,” vol. 429, pp. 1334–1353, 2008.
13. R. Koenig and J. A. Smolin, “How to efficiently select an arbitrary Clifford group element,” *J. Math. Phys.*, vol. 55, no. 12, p. 122202, Dec 2014.
14. N. Rengaswamy, R. Calderbank, S. Kadhe, and H. D. Pfister, “Logical Clifford synthesis for stabilizer codes,” *IEEE Trans. Quantum Engg.*, vol. 1, 2020. [Online]. Available: <http://arxiv.org/abs/1907.00310>
15. F. G. Brandão and M. Horodecki, “Exponential quantum speed-ups are generic,” *Quantum Information and Computation*, vol. 13, no. 11–12, pp. 901–924, 2013. [Online]. Available: <https://arxiv.org/abs/1010.3654>
16. R. Kueng, H. Zhu, and D. Gross, “Low rank matrix recovery from Clifford orbits,” *arXiv preprint arXiv:1610.08070*, 2016, [Online]. Available: <https://arxiv.org/abs/1610.08070>
17. —, “Distinguishing quantum states using clifford orbits,” *arXiv preprint arXiv:1609.08595*, 2016. [Online]. Available: <https://arxiv.org/abs/1609.08595>
18. N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, “Experimental comparison of two quantum computing architectures,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3305–3310, 2017. [Online]. Available: <https://www.pnas.org/content/114/13/3305/>
19. N. Rengaswamy, “Classical coding approaches to quantum applications,” Ph.D. dissertation, Duke University, 2020. [Online]. Available: <http://arxiv.org/abs/2004.06834>
20. R. J. McEliece, *Finite Fields for Computer Scientists and Engineers*. Kluwer Academic Publishers, 1987.
21. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.
22. R. Calderbank, E. Rains, P. Shor, and N. Sloane, “Quantum error correction via codes over GF(4),” vol. 44, no. 4, pp. 1369–1387, Jul 1998. [Online]. Available: <https://arxiv.org/abs/quant-ph/9608006>
23. N. Rengaswamy, R. Calderbank, S. Kadhe, and H. D. Pfister, “Synthesis of Logical Clifford Operators via Symplectic Geometry,” Jun 2018, pp. 791–795. [Online]. Available: <http://arxiv.org/abs/1803.06987v1>
24. D. Gottesman, “An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation,” *arXiv preprint arXiv:0904.2557*, 2009. [Online]. Available: <http://arxiv.org/abs/0904.2557>
25. A. R. Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, “The  $\mathbb{Z}_4$ -Linearity of Kerdock, Preparata, Goethals, and Related Codes,” vol. 40, no. 2, pp. 301–319, 1994.
26. P. J. Cameron and J. H. Van Lint, *Designs, Graphs, Codes and their Links*. Cambridge University Press, 1991, vol. 3.
27. Hoi Fung Chau, “Unconditionally secure key distribution in higher dimensions by depolarization,” *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1451–1468, 2005.
28. A. W. Harrow and R. A. Low, “Random quantum circuits are approximate 2-designs,” *Commun. in Math. Phys.*, vol. 291, no. 1, pp. 257–302, 2009.
29. M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, 2017.

- 
30. T. Singal and M.-H. Hsieh, "A possible improvement to the convergence analysis of the transvection markov process," personal communication, April 2021.
  31. H. Zhu, "Multiqubit clifford groups are unitary 3-designs," *Phys. Rev. A*, vol. 96, p. 062336, Dec 2017. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.96.062336>