Fermionic Averaged Circuit Eigenvalue Sampling

Adrian Chapman[1] and Steven T. Flammia[1,2]

[1]Phasecraft Inc., Washington DC, USA
[2]Department of Computer Science, Virginia Tech, Alexandria, VA, USA

arXiv:2504.01936v2 [quant-ph] 26 Mar 2026

Fermionic averaged circuit eigenvalue sampling (FACES) is a protocol to simultaneously learn the averaged error rates of many fermionic linear optical (FLO) gates simultaneously and self-consistently from a suitable collection of FLO circuits. It is highly flexible, allowing for the in situ characterization of FLO-averaged gate-dependent noise under natural assumptions on a family of continuously parameterized one- and two-qubit gates. We rigorously show that our protocol has an efficient sampling complexity, owing in-part to useful properties of the Kravchuk transformations that feature in our analysis. We support our conclusions with numerical results. As FLO circuits become universal with access to certain resource states, we expect our results to inform noise characterization and error mitigation techniques on universal quantum computing architectures which naturally admit a fermionic description.

1 Introduction

As fundamental building blocks of our universe, fermions are naturally the mediators of quantum information in many quantum computing architectures, such as superconducting [1] and semiconductor devices [2]. Many of the systems we hope to simulate with a quantum computer are systems of interacting fermions. These include quantum materials and quantum chemical systems, where the relevant degrees of freedom are interacting electrons [3]. It is thus natural to consider models of quantum computation that regard fermions as central objects. In principle, this removes any overhead from switching between fermion- and spin-based encodings, but it also presents unique challenges for other quantum information processing tasks due to the nonlocal nature of fermionic antisymmetric exchange statistics.

Predominant among these tasks is the characterization of quantum noise. Whereas noise is usually assumed to act independently on local degrees of freedom, subsequent gates will cause the effect of any such noise process to spread. In particular, fermionic models of quantum computing generally involve families of continuously parameterized gates. In this setting, even local noise will become complex—and thus difficult to learn—on the circuit timescales necessarily to realize quantum computational advantage. To address this issue, we can take inspiration from solutions in the qubit setting. In particular, we consider the strategy of *noise tailoring*, whereby noise is first physically driven to a desired form via averaging, or *twirling*, and the averaged noise parameters can be more easily learned [4, 5].

Noise tailoring is well-suited to the setting of gate-dependent noise, where we assume each distinct physical operation on the experiment introduces a unique, reproducible noise channel. The general problem of estimating all noise in a gate set having gate-dependent noise *without* any averaging is quite challenging [6, 7, 8], so it is natural to ask what

role noise tailoring can play to simplify the problem. Ref. [9] introduced averaged circuit eigenvalue sampling (ACES) for simultaneously characterizing the gate-dependent noise of many noisy Clifford gates with Pauli twirling. Clifford circuits with Pauli noise preserve the set of Pauli observables up to an overall factor—called the circuit eigenvalue—which quantifies the noise in the circuit. The main idea behind ACES is that the eigenvalues of many different noisy circuits are all dependent on a common set of gate-level noise eigenvalues, which can be extracted with access to enough experimental data.

It is well-understood that Clifford circuits with stabilizer measurements are not universal for quantum computation. Any attempt to extend the ACES protocol to a universal circuit family will thus require a generalization to circuits and noise models that do not preserve unnormalized Pauli operators. In this work, we provide a protocol for the characterization of noisy *fermionic linear optical* (FLO) circuits, or equivalently, their qubit-embedded versions, noisy nearest-neighbor *matchgate* circuits. We call this FACES, for *fermionic averaged circuit eigenvalue sampling*. FLO circuits describe general time-dependent free-fermion dynamics, and matchgate circuits describe these same dynamics on qubits under a Jordan-Wigner transformation [10]. We work almost exclusively in the fermion picture here, but our results are interchangable across the two paradigms.

Though FLO and matchgate circuits are comprised of non-Clifford gates, they are nonetheless efficiently simulable classically. However, they become universal for quantum computation when we additionally include the application of arbitrary Clifford gates; for example, even adding nearest-neighbor SWAP operations between qubits is sufficient. Similarly, Clifford circuits become universal for quantum computation with the inclusion of arbitrary matchgates. Since characterizing gate-dependent noise in Clifford circuits is a well-studied problem, we expect that FACES can be used in tandem with other methods to estimate averaged noise in universal gate sets. Furthermore, we expect that FACES will find natural applications in the setting of quantum materials and chemistry simulations, for which fermionic degrees of freedom are fundamental.

Our main results include an algorithm for extracting the FLO-averaged eigenvalues of the elements of an elementary set of FLO unitary gates. Additionally, we rigorously prove that our protocol has an efficient sample complexity when reasonable assumptions about our setting are met. Finally, we demonstrate the efficacy of our protocol by numerically showing that it is feasible for realistic error rates.

The input to our protocol is a model, which we define as follows

Definition 1 (Model)**.** *A FACES model on $2n$ fermionic modes is specified by a 3-tuple*

$$\mathcal{M} = (\{G_1, \ldots, G_K\}, \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K\}, \{C_1, \ldots, C_J\}) \tag{1}$$

for K gates G_k, each of which is associated to a vector of $2n+1$ channel eigenvalues $\boldsymbol{\xi}_k$, and $J \geq K$ FLO circuits C_j. From these, we compute the design matrix $\mathbf{A}$, whose integer-valued element A_{jk} specifies the number of occurrences of gate G_k in circuit C_j. A FACES experiment constructs this matrix and samples each circuit S times each. We say that a model is realizable by our experiment if the noisy gates are described by a set of true values for $\boldsymbol{\xi} \coloneqq \cup_k \boldsymbol{\xi}_k$.

Our main result applies to the realizable setting (i.e. we don't cover the *agnostic setting*, where we allow for the possibility that the true gate noise is not represented by the model). It is stated informally as follows.

Result 2 (Informal Statement of theorem 13)**.** *Consider the FACES model $\mathcal{M} = (\mathcal{G}, \boldsymbol{\xi}, \mathcal{C})$. A FACES experiment on $\mathcal{M}$ returns an estimate $\hat{\boldsymbol{\xi}}$ for the true gate eigenvalues $\boldsymbol{\xi}$ satisfying*

$$\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_\infty \leq \|\mathbf{A}^+\|_\infty \, \varepsilon, \tag{2}$$

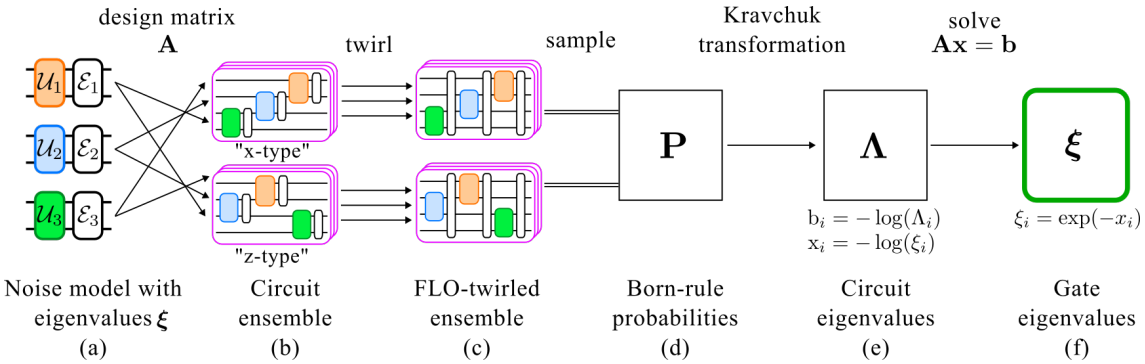


Figure 1: A graphical outline of the FACES protocol. (a) Given a set of noisy gates $\{\mathcal{U}_i\mathcal{E}_i\}_i$ we wish to characterize, we model each such gate as the associated noise channel $\mathcal{E}_i$, followed by an application of the ideal gate $\mathcal{U}_i$. (b) We compose the gates into a collection of circuits according to the design matrix $\mathbf{A}$. (c) For each circuit $\mathcal{C}$ in the collection, we randomly sample from an ensemble of circuits whose average is $\mathcal{C}$ subject to matchgate-twirled noise. (d) We sample from the circuits in our collection to obtain an estimate for the Born-rule probabilities of the associated outcomes. (e) We apply a Kravchuk transformation to obtain the eigenvalues of each circuit. (f) We fit the gate eigenvalues to the data using a log-linear model.

where $\mathbf{A}^+$ *is the pseudoinverse of the design matrix, with* $S = O(\frac{n\log(J)}{\varepsilon^2})$ *samples per circuit.*

The norm on $\mathbf{A}^+$ is the induced $\infty \to \infty$ norm, defined as $\|X\|_\infty = \max_i \sum_j |X_{ij}|$. In the setting where $\mathbf{A}$ does not have full column-rank, its psuedoinverse is still well-defined, but the model includes at least one unlearnable, or "gauge" parameter that is inaccessible to our framework (for a detailed investigation of this, see [11, 12]).

Considering eq. (2), one imagines minimizing $\|\mathbf{A}^+\|_\infty$ as a promising method to likely improve the precision of the estimate. Previous work has considered optimizing gate-dependent noise learning in a restricted setting (this is done in gate-set tomography by amplifying certain errors using long sequences [8], for example). Other methods have also been explored in an ACES-type context. For example, the authors of ref. [13] consider optimizing a figure of merit related to the gate-eigenvalue estimator covariance matrix. A comprehensive theory is lacking and currently constitutes an open research direction.

The characterization of noisy fermionic linear optical gates is well-studied in the literature [14, 15, 16, 17]. Of particular relevance to our works are the results of refs. [18] and [19]. In ref. [18], the authors propose a method for estimating the average fidelity of noisy matchgate circuits. In [19], the authors introduce a corresponding protocol that does not require an average over the matchgate group. Our result extends the former result to the in situ case under similar physical requirements, though it still requires a group average. Specifically, we require the ability to randomly sample unitaries to perform the FLO twirling and that the effect of the noise on the twirling gates can itself be incorporated into our framework. We return to this point throughout the paper.

We also require that we are able to reliably prepare the all-zeros state $|\mathbf{0}\rangle$ and all-plus state $|+\rangle^{\otimes n}$, and that we generate a set of noisy FLO circuits that act as a particular target unitary on-net (either the identity, or a particular unitary U_+, to be defined). We furthermore need to be able to measure in the computational basis and the Pauli Y-basis on at least one qubit.

Our paper is organized as follows: in section 2, we detail the relevant background on fermionic linear optical circuits (section 2.1), Kravchuk matrices (section 2.2), twirling (section 2.3), and Pauli channel eigenvalues (section 2.4). In section 3, we explain our

measurement protocol, including our assumptions (section 3.1 and section 3.2), stages of the protocol (section 3.3) and sample complexity (section 3.4). In particular, we include a proof of our sample complexity guarantee in the main text. Finally, we provide a numerical demonstration of our protocol in section 4 and conclude in section 5. Our protocol in pseudocode format is detailed in section A, and remaining proofs are given in section B.

2 Background and Notation

2.1 Fermionic Linear Optics

Fermionic linear optical (FLO) circuits constitute the most general group of unitary circuits that describe free-fermion dynamics. This group is isomorphic to the group of matchgate circuits on many-body qubit systems by the Jordan-Wigner transformation. For our purposes, we mainly keep the discussion phrased in terms of fermionic modes, with the understanding that all of our results admit corresponding descriptions in terms of qubits.

We consider FLO circuits acting on even-many fermion modes. For such a system of $2n$ modes, define the set of Majorana operators satisfying the canonical anticommutation relations

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} I, \quad \text{and} \quad \gamma_\mu^\dagger = \gamma_\mu \tag{3}$$

for all μ, $\nu \in [2n]$, where the notation $[n]$ means the set of integers $\{1, \ldots, n\}$. Denote the associated group of FLO circuits as $\mathbb{F}(2n)$. The elements of this group are Gaussian unitaries up to reflection $U \coloneqq \gamma_1^x \exp(\sum_{\mu,\nu} h_{\mu\nu}\gamma_\mu\gamma_\nu)$ for $x \in \{0, 1\}$ and $\mathbf{h}$ a real antisymmetric matrix. These preserve linear combinations of Majorana modes under conjugation. Writing $\mathcal{U}(A) \equiv UAU^\dagger$ as the unitary conjugation channel by an element $U \in \mathbb{F}(2n)$, we have

$$\mathcal{U}(\gamma_\mu) = \sum_{\nu=1}^{2n} R_{\mu\nu}\gamma_\nu \tag{4}$$

for $\mathbf{R} = \mathbf{R}_1^x e^{-4\mathbf{h}} \in O(2n)$ the single-particle transition matrix and $\mathbf{R}_1$ the canonical reflection with elements $(\mathbf{R}_1)_{\mu\nu} = -(-1)^{\delta_{\mu,1}}\delta_{\mu\nu}$. By the canonical anticommutation relations, eq. (3), FLO circuits also preserve fixed-degree polynomials of Majorana operators. For a subset $\boldsymbol{\alpha} \subseteq [2n]$, let $\gamma_{\boldsymbol{\alpha}} \equiv \prod_{\mu\in\boldsymbol{\alpha}} \gamma_\mu$ denote the *Majorana monomial* labeled by $\boldsymbol{\alpha}$, where the order of the product is taken such that the index is ascending from left to right. Eq. (4) generalizes as

$$\mathcal{U}(\gamma_{\boldsymbol{\alpha}}) = \sum_{\{\boldsymbol{\beta}\subseteq[2n]||\boldsymbol{\beta}|=|\boldsymbol{\alpha}|\}} \det(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}})\gamma_{\boldsymbol{\beta}} \tag{5}$$

where $\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ denotes the submatrix of $\mathbf{R}$ with rows taken from $\boldsymbol{\alpha}$ and columns taken from $\boldsymbol{\beta}$ (that is, Majorana monomials transform by a *compound* of $\mathbf{R}$). We adopt the notational shorthand

$$\boldsymbol{\alpha}|k, \ell \to \{\boldsymbol{\alpha} \subseteq [\ell]||\boldsymbol{\alpha}| = k\}\,. \tag{6}$$

Linear transformations of the form in eq. (5) constitute a representation of $O(2n)$ by the Cauchy-Binet formula, and this representation is irreducible. As explained in ref. [18], FLO transformations are uniquely determined up to an overall phase by their corresponding element of $O(2n)$.

It is convenient to generalize our index-subset notation in eq. (6) to binary strings with a fixed Hamming weight. We refer to a binary string by a bold alphabetic character, and, using the same notation as in eq. (6), we denote $\mathbf{x}|k, \ell \to \{\mathbf{x} \subseteq \mathbb{Z}_2^\ell || \mathbf{x}| = k\}$. In both use cases, we drop explicit reference to ℓ where clear from context and simply write "$\boldsymbol{\alpha}|k$" or "$\mathbf{x}|k$".

We also need to consider these objects in terms of Pauli operators. Denote by $\mathbb{P}(n)$ the n-qubit Pauli group, and let $\mathsf{P}(n)$ denote this group modulo an overall phase. We can label an n-qubit Pauli operator (with a convenient choice of phase) by a $2n$-bit string $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ as $\sigma_{\mathbf{x}} = i^{\mathbf{a}\cdot\mathbf{b}} \otimes_{j=1}^n X_j^{a_j} Z_j^{b_j}$, denoting single-qubit Pauli operators on a fixed site by $\{X_j, Y_j, Z_j\}$. Since we are exclusively concerned with Paulis modulo phase, multiplication in $\mathbb{P}(n)$ maps conveniently to mod-2 addition of the associated binary strings in $\mathsf{P}(n)$.

As stated earlier, Majorana monomials and Pauli operators are in one-to-one correspondence (up to a phase) by the Jordan-Wigner transformation

$$\begin{cases} \gamma_{2j-1} = Z^{\otimes(j-1)} \otimes X_j \\ \gamma_{2j} = Z^{\otimes(j-1)} \otimes Y_j \end{cases} . \tag{7}$$

These operators satisfy the canonical anticommutation relations eq. (3). In this context, the group of FLO circuits is isomorphic to the group of matchgate circuits, which we denote as $\mathbb{M}(n)$. Matchgate circuits can be decomposed into 2-qubit matchgates of the form

$$G_j(A, B) \coloneqq \begin{pmatrix} a_{00} & 0 & 0 & a_{01} \\ 0 & b_{00} & b_{01} & 0 \\ 0 & b_{10} & b_{11} & 0 \\ a_{10} & 0 & 0 & a_{11} \end{pmatrix} \tag{8}$$

on nearest-neighbor qubits $\{j, j+1\}$ on a 1-d line. Namely, A acts on the even-parity 2-qubit subspace of these qubits, and B on the odd-parity subspace. We require $\det(A) = \det(B)$ for $G_j(A, B)$ to be a matchgate. Note that, in particular, single-qubit rotations of the form $e^{i\theta Z_j}$ are matchgates for all $\theta \in [0, 2\pi)$ and all $j \in [n]$. Where clear from context, we refer to the groups $\mathbb{F}(2n)$ and $\mathbb{M}(n)$ interchangeably. For example, by $\mathcal{U}(\sigma_{\mathbf{x}})$ for $\mathcal{U} \in \mathbb{F}(2n)$, we mean conjugation by the matchgate circuit associated to $\mathcal{U}$ under the Jordan-Wigner transformation.

2.2 Kravchuk Matrices

A central object for our analysis is the following matrix

Definition 3. *(Kravchuk matrix) The* Kravchuk matrix *of order ℓ, denoted by $\mathbf{M}^{(\ell)} \in \mathbb{Z}^{(\ell+1)\times(\ell+1)}$, is an integer-valued matrix whose elements are symmetric polynomials evaluated on ± 1.*

$$M_{jk}^{(\ell)} \coloneqq e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}), \tag{9}$$

where $j,\ k \in [\ell] \cup \{0\}$, $e_k(\boldsymbol{v})$ is the degree-k elementary symmetric polynomial in the elements of the array $\boldsymbol{v}$, and $\mathbf{1}_j$ is the length-j vector of all ones.

The Walsh-Hadamard matrix of size 2^n has matrix elements $(-1)^{\mathbf{x}\cdot\mathbf{y}}$ in the standard basis, where $\mathbf{x}$ and $\mathbf{y}$ are bit-strings of length n that label matrix elements. The elements of the Kravchuk matrix are obtained by summing $\sum_{\mathbf{y}|k}(-1)^{\mathbf{x}\cdot\mathbf{y}}$ to get a matrix element that depends only on the Hamming weight k and the bit-string $\mathbf{x}$. In fact, this expression

depends only on the Hamming weight of $\mathbf{x}$. By keeping one copy of the resulting sum for each value that $|\mathbf{x}|$ can take, we get the Kravchuk matrix elements in eq. (9).

The Kravchuk matrix plays an analogous role in FACES as the Walsh-Hadamard matrix does in the original (Clifford) ACES protocol. Namely, it is the essential link between the eigenvalues and error probabilities of a FLO-twirled channel, as explained below.

We now give some alternative expressions for $\mathbf{M}^{(\ell)}$ which clearly show that this matrix can be computed efficiently. Let us introduce the notation $\mathrm{coeff}_k(p(x))$ for the coefficient of the degree-k monomial term of the polynomial $p(x)$, where the variable x will be understood from context. Then we have the following lemma.

Lemma 4 (Kravchuk matrix relations)**.** *We have the alternate expressions for* $M^{(\ell)}_{jk}$

$$M^{(\ell)}_{jk} = \mathrm{coeff}_k((1-u)^j(1+u)^{\ell-j}) \tag{10}$$

$$= \sum_{\boldsymbol{\alpha}|k,\ell} (-1)^{|\boldsymbol{\alpha}\cap[j]|} \tag{11}$$

$$M^{(\ell)}_{jk} = \sum_{m=\max(0,k+j-\ell)}^{\min(j,k)} \binom{\ell-j}{k-m}\binom{j}{m}(-1)^m . \tag{12}$$

Additionally, $\mathbf{M}^{(\ell)}$ *is proportional to an involution*

$$(\mathbf{M}^{(\ell)})^2 = 2^\ell \mathbf{I}. \tag{13}$$

Proof. The proof is the content of section B.1. □

While a proof of the above relations can be found in existing literature, we include it here in the interest of keeping our results self-contained. Where clear from context, we drop the label ℓ from the Kravchuk matrix and any related matrices.

In some cases, we need to account for phases that accumulate from commuting Majorana operators past one another. To this end, it is convenient to define a relative of the Kravchuk matrix as follows.

Definition 5. *(Antipode involution) The* antipode *involution matrix,* $\mathbf{s}^{(\ell)}$ *for* ℓ *even, is the permutation matrix with components*

$$s^{(\ell)}_{jk} = \delta_{jk}\delta_{j \pmod 2,0} + \delta_{\ell-j,k}\delta_{j \pmod 2,1}. \tag{14}$$

The antipode involution takes indices with fixed odd weight and exchanges $j \leftrightarrow \ell - j$. It satisfies the following relations due to the particular symmetry properties of $\mathbf{M}$

Lemma 6. *(Kravchuk-antipode relations) The Kravchuk matrix and antipode involution commute, and their product has components*

$$(\mathbf{s}^{(\ell)} \cdot \mathbf{M}^{(\ell)})_{jk} = (\mathbf{M}^{(\ell)} \cdot \mathbf{s}^{(\ell)})_{jk} = (-1)^{jk} e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}). \tag{15}$$

Proof. The proof is the content of section B.2. □

Finally, we define the diagonal matrix of binomial coefficients $\mathbf{d}^{(\ell)}$ by

$$d^{(\ell)}_{jk} = \binom{\ell}{k}\delta_{jk}. \tag{16}$$

It is straightforward to see that $\mathbf{s}^{(\ell)}$ and $\mathbf{d}^{(\ell)}$ commute as well, but $\mathbf{M}^{(\ell)}$ and $\mathbf{d}^{(\ell)}$ do not commute. Each of these transformations will be important for our estimation procedure.

2.3 Noise Channels and Twirling

A quantum noise channel is described by a completely positive trace-preserving (CPTP) map, admitting a Kraus representation

$$\mathcal{E}(X) = \sum_j K_j X K_j^\dagger, \tag{17}$$

where $\{K_j\}$ are the Kraus operators satisfying

$$\sum_j K_j^\dagger K_j = I. \tag{18}$$

For a unitary ensemble $\mathcal{S}$ (either discrete or continuous), we define the $\mathcal{S}$-twirling of $\mathcal{E}$ as

$$\mathcal{E}^{\mathcal{S}} = \mathbb{E}_{\mathcal{U}\in\mathcal{S}}[\mathcal{U}\circ\mathcal{E}\circ\mathcal{U}^\dagger] \tag{19}$$

where we employ the unitary channel notation from section 2.1, and $\mathbb{E}_x[f(x)]$ denotes the expectation value over x of $f(x)$. The main examples we consider are Pauli and FLO twirling, where the expectation is given respectively by

$$\mathcal{E}^{\mathbb{P}(n)}(X) = 4^{-n} \sum_{\mathbf{x}\in\mathsf{P}(n)} \sum_j (\sigma_{\mathbf{x}} K_j \sigma_{\mathbf{x}})\, X\, (\sigma_{\mathbf{x}} K_j \sigma_{\mathbf{x}})^\dagger \,, \tag{20}$$

and

$$\mathcal{E}^{\mathbb{F}(2n)}(X) = \int_{U\in\mathbb{F}(2n)} dU \;\sum_j \mathcal{U}(K_j)\, X\, \mathcal{U}(K_j)^\dagger \tag{21}$$

where eq. (21) is a Haar average over the group of FLO unitaries. Note also that twirling is idempotent. That is, twirling twice is the same as twirling once. Additionally, $\mathcal{E}^{\mathcal{S}}$ commutes with every element of $\mathcal{S}$ when $\mathcal{S}$ is a group, by construction.

A channel that is invariant under Pauli twirling is called a Pauli channel. It is easy to show that Pauli twirling a general channel results in a Pauli channel [20, Lemma 5.2.4]. Pauli channels can be parameterized as

$$\mathcal{E}^{\mathbb{P}(n)}(X) = \sum_{\mathbf{x}\in\mathsf{P}(n)} p_{\mathbf{x}} \sigma_{\mathbf{x}} X \sigma_{\mathbf{x}} \tag{22}$$

where the $\{p_{\mathbf{x}}\}_{\mathbf{x}\in\mathsf{P}(n)}$ form a probability distribution over bit-strings of length $2n$. These are called the *Pauli error probabilities*. Starting from a general channel with Kraus operators $K_j \equiv \sum_j c_{j\mathbf{x}}\sigma_{\mathbf{x}}$ as in eq. (17), the relationship to the Pauli error probabilities of the twirled channel in eq. (22) is [20, Lemma 5.2.4]

$$p_{\mathbf{x}} = \sum_j |c_{j\mathbf{x}}|^2 \,. \tag{23}$$

Similarly, a *FLO-twirled channel* is one that is invariant under FLO twirling. Since all Pauli operators are matchgates, any FLO-twirled channel is also invariant under Pauli twirling, and is therefore a Pauli channel. Note the following contrast between Pauli twirling and FLO twirling: Pauli twirling a unitary Pauli (i.e. rank-1) channel leaves that channel invariant, but it is *not* the case that FLO unitary channels are invariant under FLO twirling since they generally become non-unitary after twirling. FLO-twirled channels on $2n$ modes can be parameterized as

$$\mathcal{E}^{\mathbb{F}(2n)}(X) = \sum_{k=0}^{2n} \binom{2n}{k}^{-1} q_k \sum_{\boldsymbol{\alpha}|k,2n} \gamma_{\boldsymbol{\alpha}} X \gamma_{\boldsymbol{\alpha}}^\dagger \,, \tag{24}$$

where q_k form a probability distribution over $\{0,\ldots,2n\}$. We call the q_k the *fermionic error probabilities*.

Lemma 7 (FLO twirl of a Pauli channel)**.** *Consider a Pauli channel $\mathcal{E}^{\mathbb{P}(n)}$ with Pauli error probabilities $p_{\mathbf{x}}$. Let $\mathbf{x}(\boldsymbol{\alpha})$ denote the Pauli string associated to the subset of Majorana modes $\boldsymbol{\alpha}$ by the Jordan-Wigner transformation. Then $\mathcal{E}^{\mathbb{F}(2n)} = (\mathcal{E}^{\mathbb{P}(n)})^{\mathbb{F}(2n)}$, and the fermionic error probabilities q_k of $\mathcal{E}^{\mathbb{F}(2n)}$ are given by*

$$q_k = \sum_{\boldsymbol{\alpha}|k} p_{\mathbf{x}(\boldsymbol{\alpha})} \,. \tag{25}$$

Proof. See section B.3. □

We remark that the FLO twirl of a *general* channel has fermionic error probabilities that are related to the Kraus operators by concatenating the expressions in eqs. (23) and (25).

We model a noisy implementation $\widetilde{\mathcal{V}}$ of a FLO unitary $\mathcal{V}$ as a noise map $\mathcal{E}$ followed by the ideal channel $\mathcal{V}$, with the intuition that $\mathcal{E}$ is close to the identity channel. This provides the following

Definition 8. *(Unitary channel noise) For a noisy implementation $\widetilde{\mathcal{V}}$ of a unitary conjugation channel $\mathcal{V}$, we define the noise map corresponding to $\widetilde{\mathcal{V}}$ by*

$$\mathcal{E} := \mathcal{V}^\dagger \widetilde{\mathcal{V}}, \tag{26}$$

and this gives $\widetilde{\mathcal{V}} = \mathcal{V}\mathcal{E}$.

We will eventually require that $\mathcal{E}$, so defined, is not too noisy in a sense made precise in section 3.1. Prepending a random FLO unitary $\mathcal{U}^\dagger$ to $\widetilde{\mathcal{V}}$ and appending the unitary $\mathcal{U}$ conjugated by $\mathcal{V}$ gives the twirled noise map followed by the ideal channel in expectation. The following statement of this is a standard result.

Lemma 9 (Noise Twirling)**.** *For $\mathcal{S} \in \{\mathbb{P}(n), \mathbb{F}(2n)\}$, we have*

$$\widetilde{\mathcal{V}}^{\mathcal{S}} := \int_{\mathcal{S}} d\mathcal{U} \left(\mathcal{V}\mathcal{U}\mathcal{V}^\dagger\right)(\mathcal{V}\mathcal{E})\mathcal{U}^\dagger = \mathcal{V}\mathcal{E}^{\mathcal{S}} \,. \tag{27}$$

Proof. Immediate from the definition of twirling. □

In practice, twirling is done by taking a sample average over the ensemble of unitary circuits appearing in the twirl [4, 5]. In the Pauli-twirl case, the overhead that this induces is negligible in the sense that the depth-one Pauli unitary circuits only slightly increase the depth overall, or can often be compiled into the underlying circuit [5]. However, the best-known implementations of a generic FLO circuit require $\mathcal{O}(n)$ depth [21]. It would be interesting to investigate reducing this depth penalty by sampling generators of the FLO group and analyzing the mixing time of the resulting random walk on $O(2n)$, similar to what was done in ref. [22] for finite groups, but this is beyond the scope of the present work.

2.4 Channel Eigenvalues

If $\mathcal{E}^{\mathbb{P}(n)}$ is a Pauli channel, then $\mathcal{E}^{\mathbb{P}(n)}$ satisfies an eigenvector relation with respect to the Pauli operators

$$\mathcal{E}^{\mathbb{P}(n)}(\sigma_{\mathbf{x}}) = \lambda_{\mathbf{x}}\sigma_{\mathbf{x}} \,. \tag{28}$$

The $\lambda_{\mathbf{x}}$ are called the *Pauli channel eigenvalues*. They are related to the Pauli error probabilities by

$$\lambda_{\mathbf{x}} = \sum_{\mathbf{y}\in\mathsf{P}(n)} (-1)^{\langle \mathbf{x},\mathbf{y}\rangle} p_{\mathbf{y}} \tag{29}$$

where $\langle \mathbf{x}, \mathbf{y}\rangle$ is the symplectic inner product between bit strings $\mathbf{x}$, $\mathbf{y} \in \{0,1\}^{\times 2n}$. Up to a column permutation, this is the Walsh-Hadamard transformation.

Now let $\mathcal{E}^{\mathbb{F}(2n)}$ be a FLO-twirled channel. Then there is a similar eigenvalue relationship where the eigenvectors are Majorana monomials. From the parameterization in eq. (24), we can derive the eigenvalue relation

$$\begin{aligned}\mathcal{E}^{\mathbb{F}(2n)}(\gamma_{\boldsymbol{\beta}}) &= \sum_{k=0}^{2n} \binom{2n}{k}^{-1} q_k \left(\sum_{\boldsymbol{\alpha}|k} \gamma_{\boldsymbol{\alpha}}^\dagger \gamma_{\boldsymbol{\beta}} \gamma_{\boldsymbol{\alpha}}\right) \\ &= \left\{\sum_{k=0}^{2n} \left[\sum_{\boldsymbol{\alpha}|k} (-1)^{|\boldsymbol{\beta}|k+|\boldsymbol{\alpha}\cap\boldsymbol{\beta}|}\right] \binom{2n}{k}^{-1} q_k\right\} \gamma_{\boldsymbol{\beta}}\,.\end{aligned} \tag{30}$$

The coefficient in braces can be simplified using theorem 4 and the definitions in section 2.2 to give the *FLO-twirled channel eigenvalue*

$$\xi_j = \sum_{k=0}^{2n} (-1)^{jk} e_k(-\mathbf{1}_j \oplus \mathbf{1}_{2n-j}) \binom{2n}{k}^{-1} q_k \tag{31}$$

$$= \left[(\mathbf{s}\cdot\mathbf{M}\cdot\mathbf{d}^{-1})\cdot\mathbf{q}\right]_j . \tag{32}$$

This constitutes a proof of the following lemma, for which we adopt a vector notation.

Lemma 10. *(FLO-twirled Channel Eigenvalues) We have*

$$\boldsymbol{\xi} = \left(\mathbf{s}\cdot\mathbf{M}\cdot\mathbf{d}^{-1}\right)\cdot\mathbf{q}\,. \tag{33}$$

Proof. Immediate. □

As $2^{-\ell/2}\left[\mathbf{s}^{(\ell)}\cdot\mathbf{M}^{(\ell)}\right]$ is an involution, we can invert this relation and express the probabilities $\mathbf{q}$ from eq. (33) in terms of the eigenvalues $\boldsymbol{\xi}$ as

$$\mathbf{q} = 4^{-n}\left(\mathbf{d}\cdot\mathbf{M}\cdot\mathbf{s}\right)\cdot\boldsymbol{\xi}\,. \tag{34}$$

Owing to this linear relation, we can learn either $\mathbf{q}$ or $\boldsymbol{\xi}$ to completely characterize a FLO-twirled channel, but one must be careful about the precision in these different cases due to the relative normalization.

A FLO-twirled channel commutes with every element of the group $\mathbb{F}(2n)$, so we can extend the notion of a channel eigenvalue to circuits as

$$\widetilde{\mathcal{U}}^{\mathbb{F}(2n)} \coloneqq \prod_g \left(\mathcal{U}_g \mathcal{E}_g^{\mathbb{F}(2n)}\right) = \left(\prod_g \mathcal{U}_g\right)\left(\prod_g \mathcal{E}_g^{\mathbb{F}(2n)}\right). \tag{35}$$

This gives

$$\widetilde{\mathcal{U}}^{\mathbb{F}(2n)}(\gamma_{\boldsymbol{\beta}}) = \left(\prod_g \xi_{g,k}\right)\mathcal{U}(\gamma_{\boldsymbol{\beta}}) \qquad (k \coloneqq |\boldsymbol{\beta}|) \tag{36}$$

where $\xi_{g,k}$ is the degree-k *gate eigenvalue* for the circuit factor g. We define the product in parentheses as the degree-k circuit eigenvalue for $\widetilde{\mathcal{U}}^{\mathbb{F}(2n)}$ as

$$\Lambda_k \coloneqq \prod_g \xi_{g,k} \tag{37}$$

3 Measurement Protocol

Efficiently estimating noise in a quantum system requires that we fit to models with certain simplifying assumptions. Following from theorem 8, we assume the parameters describing the noise channel $\mathcal{E}$ are independent of time and, furthermore, depend only on the gate we apply. This is similar to the time-stationary and Markovian assumptions of [23, def. 2], but unlike ref. [23], and much related literature, we allow wide latitude for gate dependence. In that regard our protocol can be viewed as a kind of "fermionic gate-set tomography" in the FLO-twirled subspace [6, 24, 7, 8].

We broadly refer to the map that takes ideal physical operations—state preparations, unitary gate operations, and measurements—to their parameterized noisy versions as a *model* (and we view the FACES model given in theorem 1 as a special case of this). Denote the set of idealized physical operations by $\mathcal{S}$ and the set of noisy operations by $\mathcal{N}$, then the model is a bijective map $\phi : \mathcal{S} \to \mathcal{N}$. We only consider circuits composed of gates for which the model is defined. This is the *realizable setting*. If the "true" noise belongs to the realizable setting, then we provide convergence/correctness guarantees that we can learn that model. Otherwise, one of our assumptions is violated. We next present some refining assumptions about our models in the following subsection.

3.1 Model assumptions

Denote a noisy quantum circuit by $\widetilde{\mathcal{U}} = \prod_{g=1}^{L} \widetilde{\mathcal{U}}_g$, where each noisy gate factor is given by $\widetilde{\mathcal{U}}_g := \mathcal{U}_g \mathcal{E}_g$ as in theorem 8, extended to a circuit of length L as in eq. (35). The ideal part $\mathcal{U}_g$ of each noisy gate is known from the physical operation we applied, but the accompanying noise channel $\mathcal{E}_g$ contains free parameters to be learned. In principle, this definition allows for the case where we have multiple ways of physically implementing the same idealized circuit, resulting in multiple circuits with the same ideal part but possibly distinct accompanying noise channels. As stated above, we will restrict to the setting where the applied noise channel is unique to the corresponding idealized gate (that is, our model ϕ is bijective), but it is straightforward to relax this assumption. In addition, the noisy part of each gate may also depend on the idealized gate itself. In principle, the twirling gates $\mathcal{U}^\dagger$ and $\mathcal{V}\mathcal{U}\mathcal{V}^\dagger$ in eq. (27) are noisy in the FLO-twirling case as well, but we assume that this noise is only weakly correlated with $\mathcal{V}$, since we can implement them without necessarily implementing $\mathcal{V}$ due to the FLO group structure. Our expectation is that, in practice, this noise will introduce some additional uncertainty on our gate-eigenvalue estimates. Under these assumptions, the FACES model given in theorem 1 is a special case of the general model described in this section where each individual FLO unitary gate is associated to its FLO-twirled noisy version through its collection of gate eigenvalues. In summary, we assume that any noise model we wish to learn is Markovian in at least two senses: first, that it can be described by a completely positive trace-preserving map, but also that the noise specific to a given gate depends only on the idealized gate itself.

We require that our state preparations and measurements (SPAM) are not too noisy so that we can obtain useful information from experiments. Specifically, we define $\widetilde{\rho}_{\mathbf{x}}$ to be the density matrix describing a noisy preparation of the computational basis state $|\mathbf{x}\rangle\langle\mathbf{x}|$. Similarly, let $\widetilde{E}_{\mathbf{x}}$ correspond to the noisy POVM element associated to measuring the computational basis state $|\mathbf{x}\rangle\langle\mathbf{x}|$. Our first assumption is that SPAM errors are sufficiently weak, as quantified by the following inequality,

$$\langle \mathbf{0}| \widetilde{\rho}_{\mathbf{0}} |\mathbf{0}\rangle \langle \mathbf{x}| \widetilde{E}_{\mathbf{x}} |\mathbf{x}\rangle \geq 1 - c \tag{38}$$

for some sufficiently small constant c, say $c \leq \frac{1}{3}$. The noise channels $\mathcal{E}$ we consider should be also such that their FLO-twirled versions are not too noisy, otherwise they will be hard to learn. We quantify this mathematically as

$$||\mathcal{I} - \mathcal{E}_j^{\mathbb{F}(2n)}||_\infty \leq c' \tag{39}$$

for all noise channels $\mathcal{E}_j$ that we are trying to learn. Here again, c' is a sufficiently small constant. Since our channels are always diagonal in the Pauli basis, this is equivalent to requiring that any noisy-channel eigenvalue lies in the interval $[1 - c', 1]$ (in the proof of theorem 13, we require that $c' = \frac{1}{2}$). In practice, we routinely expect that these numbers c, c' are less than 0.05 or even 0.01 in some quantum architectures, so these are not strong assumptions.

3.2 FLO-Specific Considerations

A fundamental challenge attendant to the setting of noisy FLO circuits arises due to the fact that the group $\mathbb{F}(2n)$ is continuous. Since we can only ever estimate finitely many noise parameters, we thus assume that the noise can nonetheless be described by a finite set of channels. For completeness, we choose a particular gate set given by $\mathcal{G} := \{e^{i\theta Z_j}\}_{j,\theta} \cup \{G_j(H,H)\}_j$ in the qubit picture, where $\theta \in [0, 2\pi)$ is allowed to be an arbitrary continuous parameter for each j. The matrix H is the single-qubit Hadamard operator, and $G_j(A,B)$ is the 2-qubit matchgate as defined in section 2.1. While the 2-qubit gates $\{G_j(H,H)\}_j$ form a discrete set, the single qubit rotations $\{e^{i\theta Z_j}\}_{j,\theta}$ do not. We partition the interval $[0, 2\pi)$ into N bins of the form $B_k := [\frac{2\pi(k-1)}{N}, \frac{2\pi k}{N})$ for $k \in [N]$, and we assume that the single-qubit rotation $e^{i\theta Z_j}$ is subject to a fixed noise channel $\mathcal{E}_{1,jk}$ if $\theta \in B_k$. Similarly, we assume that each of the $\{G_j(H,H)\}_j$ is subject to a fixed noise channel $\mathcal{E}_{2,j}$ An interesting related question is whether we can use these methods to learn a finite set of *noise generators* for a continuous family of noise operations. We leave this question for future work, and we return to it in the Discussion.

Finally, we review some practical details related to twirling over the FLO group. Importantly, FLO twirling cannot generally be made to be constant-depth. One way to see this is to note that, in the qubit picture, matchgate twirling averages together the error probabilities of local Paulis with those of arbitrarily high weight (e.g. X_1 and $Z^{\otimes(n-1)}Y_n$). Thus, any FLO unitary ensemble we use to apply the twirling must involve at least one deep circuit if we require our gates to have constant support. We again return to these points in the Discussion.

3.3 Protocol Details

Our measurement protocol for obtaining FLO-averaged circuit eigenvalues closely follows the ACES protocol of ref. [9], and it is detailed graphically in fig. 1. To learn all the eigenvalues of a FLO-twirled noise channel, we present two complementary protocols, which differ slightly in the required input state, measurement basis, and net FLO circuit applied. An important unitary conjugation channel in our analysis is $\mathcal{U}_+(X) = U_+ X U_+^\dagger$, related to the unitary $U(R_+)$ defined in [18], and satisfying

$$\mathcal{U}_+(|+\rangle\langle+|^{\otimes n}) = \left(\frac{I + i\gamma_1}{\sqrt{2}}\right) |\mathbf{0}\rangle\langle\mathbf{0}| \left(\frac{I - i\gamma_1}{\sqrt{2}}\right). \tag{40}$$

We have the following decomposition in terms of the gate set from the previous section

$$U_+ = e^{-i\frac{\pi}{4}Z_1} \prod_{j=1}^{n-1} \left[e^{-i\frac{\pi}{4}Z_{j+1}} G_j(H,H) e^{i\frac{\pi}{4}Z_j} G_j(H,H) e^{i\frac{\pi}{4}Z_{j+1}}\right] \tag{41}$$

up to an overall phase, where the product is taken in ascending order in j from left to right.

In general, we label an ensemble of noisy FLO circuits as

$$\mathcal{C}_r = \left\{ \widetilde{\mathcal{U}}^{(j)} \coloneqq \prod_{m=1}^{L(j)} \widetilde{\mathcal{U}}_m^{(j)} \right\}_j \tag{42}$$

where $r \in \{\text{“}x\text{”}, \text{“}z\text{”}\}$, and the net ideal unitary varies with r as follows: for all $\widetilde{\mathcal{U}}^{(j)} \in \mathcal{C}_z$, we have $\mathcal{U}^{(j)} = \mathcal{I}$ (the identity channel), and for all $\widetilde{\mathcal{U}}^{(j)} \in \mathcal{C}_x$ we have $\mathcal{U}^{(j)} = \mathcal{U}_+$. Recall that a tilde is used to indicate the noisy version of a gate or circuit. To avoid overloading the superscript notation, we denote $\mathcal{C}_r^{\mathbb{F}(2n)}$ as the ensemble whose elements correspond to the FLO-twirled elements of $\mathcal{C}_r$, and we drop the explicit superscript on the elements of $\mathcal{C}_r^{\mathbb{F}(2n)}$ themselves, with the understanding that they are twirled.

We can organize our circuits into two cases depending on whether the first qubit is prepared in the x or z basis; we call these “type-x” and “type-z”, respectively. The readout data, and thus the expected signal, takes two different forms in these two cases. We thus organize our empirical estimates of the outcome probabilities according to the circuit (or equivalently, readout) type. Readout data of type-z is the Hamming weight of an n-bit string, whereas readout data of type-x consists of a sign ($\pm$) together with the Hamming weight of an $(n-1)$-bit string. When we collect these data into empirical probability estimates over the relevant Hamming weights, we denote these empirical estimates by $\mathbf{P}$ regardless of the x/z type. Nevertheless, we can distinguish the type by adding an index s that takes values in $\{+,-,0\}$ where 0 indicates z-type data, and $\pm$ corresponds to the sign label of x-type data. For example, the estimated probability of weight-j in a $+$-outcome x-type experiment would be labeled $\mathrm{P}_{+,j}$, and so on.

We define the z-type Born-rule probability distribution sampled in experiment j as

$$\mathrm{P}_{0,\ell}^{(j)} \coloneqq \sum_{\mathbf{w}|\ell,n} \langle \mathbf{w}| \widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) |\mathbf{w}\rangle \tag{43}$$

for $\widetilde{\mathcal{U}}^{(j)} \in \mathcal{C}_z^{\mathbb{F}(2n)}$. Similarly, define the x-type Born-rule probability distribution in experiment j as

$$\mathrm{P}_{\pm,\ell}^{(j)} \coloneqq \sum_{\mathbf{w}|\ell,n-1} \langle \pm_y \mathbf{w}| \widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n}) |\pm_y, \mathbf{w}\rangle \,. \tag{44}$$

Here, the first qubit is measured in the y-basis, and the outcome of this measurement corresponds to the $\pm$ index on this probability. Let $\mathbf{P}_s^{(j)}$ be the vector with components $\mathrm{P}_{s,\ell}^{(j)}$ for $s \in \{+,-,0\}$

The Born-rule probability distributions are related to the circuit eigenvalues of a FLO-twirled quantum circuit. Let $\Lambda_k^{(j)}$ denote the degree-k eigenvalue of circuit j in either ensemble $\mathcal{C}_x^{\mathbb{F}(2n)}$ or $\mathcal{C}_z^{\mathbb{F}(2n)}$, for $k \in [2n]$. We collect these eigenvalues into vectors as

$$\mathbf{\Lambda}_{\text{even}}^{(j)} \coloneqq \left(\Lambda_{2k}^{(j)}\right)_{k=0}^{n} \tag{45}$$

$$\mathbf{\Lambda}_{\text{even}\backslash\{2n\}}^{(j)} \coloneqq \left(\Lambda_{2k}^{(j)}\right)_{k=0}^{n-1} \tag{46}$$

$$\mathbf{\Lambda}_{\text{odd}}^{(j)} \coloneqq \left(\Lambda_{2k+1}^{(j)}\right)_{k=0}^{n-1}, \tag{47}$$

and we have

Lemma 11. *For an ensemble of noisy and FLO-twirled FLO circuits* $\mathcal{C}_r^{\mathbb{F}(2n)}$,

$$\mathbf{P}_0^{(j)} = 2^{-n}\mathbf{d}^{(n)} \cdot \mathbf{M}^{(n)} \cdot \mathbf{\Lambda}_{\text{even}}^{(j)} \tag{48}$$

for $r =$ "z". *Similarly,*

$$\frac{1}{2}(\mathbf{P}_+^{(j)} + \mathbf{P}_-^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \mathbf{\Lambda}_{\text{even}\setminus\{2n\}}^{(j)} \tag{49}$$

$$\frac{1}{2}(\mathbf{P}_+^{(j)} - \mathbf{P}_-^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \mathbf{\Lambda}_{\text{odd}}^{(j)} \tag{50}$$

for $r =$ "x".

We collect these transformations into a global one over all experiments as

$$\mathbf{\Lambda} = \mathbf{V}\mathbf{P} \tag{51}$$

where

$$\mathbf{V} = \mathbf{V}_z \oplus \mathbf{V}_x \tag{52}$$

and

$$\begin{cases} \mathbf{V}_z = \bigoplus_{j\in\mathcal{C}_z} \left[\mathbf{M}^{(n)} \cdot (\mathbf{d}^{(n)})^{-1}\right] \\ \mathbf{V}_x = \bigoplus_{j\in\mathcal{C}_x} \left\{\left[\mathbf{M}^{(n-1)} \cdot (\mathbf{d}^{(n-1)})^{-1}\right] \oplus \left[\mathbf{M}^{(n-1)} \cdot (\mathbf{d}^{(n-1)})^{-1}\right]\right\} \cdot (\sqrt{2}\mathbf{H} \otimes \mathbf{I}_n) \end{cases} \tag{53}$$

with $\mathbf{H}$ is the 2×2 Hadamard transform. Using these transformations, we reconstruct the global set of circuit eigenvalues and subsequently reconstruct the the gate eigenvalues. Our protocol proceeds as follows: given a set of elementary gates $\mathcal{G}$ we would like to characterize, we implement an ensemble of these circuits composed of the corresponding FLO-twirled gates. Each circuit in our ensemble has eigenvalue given by the design matrix

$$\Lambda_k^{(j)} = \prod_{g\in\mathcal{G}} \xi_{g,k}^{A_{j,g}} \tag{54}$$

where the design matrix element $A_{j,g}$ corresponds to the number of occurences of gate g in circuit j. Physically, we sample the Born rule probabilities defined in eq. (43) and eq. (44) to obtain empirical estimates $\hat{\mathbf{P}}^{(j)}$ of these distributions for each circuit j, and we apply the transformations in theorem 11 to produce estimators $\hat{\mathbf{\Lambda}}^{(j)}$ for the associated circuit eigenvalues. Finally, we fit our gate eigenvalue estimates to our circuit eigenvalue data using a *log-linear model*. That is, our estimator for each Majorana degree k is the solution to a system of equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{A}$ is the design matrix, $x_{g,k} = -\log(\xi_{g,k})$, and $b_{j,k} = -\log(\Lambda_k^{(j)})$. More formally, we give the following definition

Definition 12 (Estimator)**.** *For a FACES model* $\mathcal{M} = (\mathcal{G}, \boldsymbol{\xi}, \mathcal{C})$*, from which we obtain the design matrix* $\mathbf{A}$*, the FACES estimator for* $\hat{\boldsymbol{\xi}}$ *is taken as* $\hat{\xi}_{g,k} = \exp(-\hat{x}_{g,k})$ *for each* g*,* k*, with*

$$\hat{x}_{g,k} = \begin{cases} \left(\mathbf{A}^+\hat{\mathbf{b}}\right)_{g,k} & \left(\mathbf{A}^+\hat{\mathbf{b}}\right)_{g,k} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{55}$$

That is, we take the estimator given by ordinary least-squares, projected onto the positive orthant for each Majorana degree k. One could of course consider other estimators, such as $\hat{\mathbf{x}}'$ taken to be the element of the positive orthant that minimizes $\|\mathbf{A}\hat{\mathbf{x}}' - \hat{\mathbf{b}}\|_2$, or even an estimator that maximizes a likelihood function subject to a positivity constraint. These are convex programs and would presumably be efficient to consider as practical estimators. However, our choice of estimator is convenient from the standpoint of proving sample-complexity guarantees. We give a pseudocode description of the relevant subroutines in section A.

3.4 Sample Complexity

In this section, we analyze the sample complexity required to bound the error on the estimator $\hat{\boldsymbol{\xi}}$ in theorem 12. We provide the following theorem:

Theorem 13. *Suppose we have both* $\Lambda_k^{(j)} \geq \frac{1}{2}$ *and* $\hat{\Lambda}_k^{(j)} \geq \frac{1}{4}$ *for all degree sectors* k *and experiments* j*, then the FACES estimator satisfies*

$$\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_\infty \leq 4\|\mathbf{A}^+\|_\infty \varepsilon \tag{56}$$

We thus require $S = O\left(\frac{n\|\mathbf{A}^+\|_\infty^2 \log(J/\delta)}{\epsilon^2}\right)$ *samples per experiment to guarantee error* ϵ *with probability* $1 - \delta$*.*

We prove this in two parts. We first prove a lemma that bounds the sensitivity of the gate-noise eigenvalue estimates in terms of the sample estimates of the Born rule probabilities of each experiment. The main theorem on sample complexity then follows from well-known results about estimating probability distributions from samples in the 1-norm.

Lemma 14. *Suppose we have both* $\Lambda_k^{(j)} \geq \frac{1}{2}$ *and* $\hat{\Lambda}_k^{(j)} \geq \frac{1}{4}$ *for all degree sectors* k *and experiments* j*, then for all* $\varepsilon \in (0, 1/4]$ *we have*

$$\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_\infty \leq 4\|\mathbf{A}^+\|_\infty \varepsilon \tag{57}$$

where

$$\max_j \|\delta \mathbf{P}^{(j)}\|_1 \leq \varepsilon \tag{58}$$

is a uniform upper bound on the 1-norm error in the estimate of the Born rule probabilities.

Here the matrix norm $\|\cdot\|_\infty$ is defined in terms of the vector ∞-norm as

$$\|\mathbf{A}\|_\infty = \max_{x \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \tag{59}$$

and can be computed as the maximum value of the vector 1-norm of each row.

Our proof relies on several applications of the mean value theorem, which we recall is stated as:

Lemma 15 (Mean-Value Theorem)**.** *Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function on the closed interval* $[a, b]$ *and differentiable on the open interval* (a, b) *for* $a < b$*. There exists some* c *in* (a, b) *such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}\,. \tag{60}$$

Proof. (theorem 14) The logarithms of the circuit and gate eigenvalues satisfy a linear relationship

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{61}$$

where $x_i = -\log(\xi_i)$, and $b_i = -\log(\Lambda_i)$, and we have bundled the compound indices into single indices for simplicity. Letting $\delta x_i \equiv \hat{x}_i - x_i$, define the function

$$u_i(t) = \exp\left(x_i + t\delta x_i\right) \tag{62}$$

for each i, such that $u_i(0) = \xi_i$ and $u_i(1) = \hat{\xi}_i$. Applying theorem 15 to each such function separately,

$$\hat{\xi}_i - \xi_i = u_i(1) - u_i(0) \tag{63}$$
$$= u_i'(c_i) \tag{64}$$
$$\hat{\xi}_i - \xi_i = \delta x_i \exp\left(x_i + c_i \delta x_i\right) \tag{65}$$

for some $c_i \in (0, 1)$ The exponential function is monotonic, so we are guaranteed that the absolute value of the error is less than or equal to the value of the right-hand side for $c_i = 0$ (if $\delta x_i < 0$) or $c_i = 1$ (if $\delta x_i > 0$). Using the definition of the gate eigenvalue and the estimate thereof gives

$$|\hat{\xi}_i - \xi_i| \leq |\delta x_i| \max(\hat{\xi}_i, \xi_i) \leq |\delta x_i| \tag{66}$$

where the second inequality follows since we are assuming the true gate eigenvalue and its estimate are upper-bounded by 1. This gives the inequality in terms of vector norms

$$\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_\infty \leq \|\delta \mathbf{x}\|_\infty. \tag{67}$$

Next, let $\tilde{\mathbf{x}} \coloneqq \mathbf{A}^+ \hat{\mathbf{b}}$ be the ordinary least-squares estimator. We have

$$|\delta x_i| = \begin{cases} |\tilde{x}_i - x_i| & \tilde{x}_i > 0 \\ |x_i| & \tilde{x}_i \leq 0 \end{cases} \tag{68}$$

for all i. If $\tilde{x_i} \leq 0$, we have

$$|x_i| \leq |x_i + |\tilde{x}_i|| = |x_i - \tilde{x}_i| = |\tilde{x}_i - x_i| \tag{69}$$

Thus $|\delta x_i| \leq |\tilde{x}_i - x_i|$ for all i, and this leads to $\|\delta \mathbf{x}\|_\infty \leq \|\tilde{\mathbf{x}} - \mathbf{x}\|_\infty$. By the assumptions that we are in the realizable setting and that $\mathbf{A}$ is full-rank, we have $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, and this gives

$$\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_\infty \leq \|\mathbf{A}^+ (\hat{\mathbf{b}} - \mathbf{b})\|_\infty \leq \|\mathbf{A}^+\|_\infty \|\hat{\mathbf{b}} - \mathbf{b}\|_\infty \tag{70}$$

where the second inequality follows from submultiplicativity of the norm.

We bound the error on $\mathbf{b}$ similarly. Letting $\delta\Lambda_i = \hat{\Lambda}_i - \Lambda_i$, we define

$$v_i(t) = \log(\Lambda_i + t\delta\Lambda_i) \tag{71}$$

for each index i such that $v_i(0) = b_i$ and $v_i(1) = \hat{b}_i$. We have

$$\hat{b}_i - b_i = v_i(1) - v_i(0) \tag{72}$$
$$= v_i'(\tilde{c}_i) \tag{73}$$
$$\hat{b}_i - b_i = \frac{\delta\Lambda_i}{\Lambda_i + \tilde{c}_i \delta\Lambda_i} \tag{74}$$

for $\tilde{c}_i \in (0, 1)$. The function on the right-hand side is monotonic in $\tilde{c}_i$, so its absolute value is maximized for $\tilde{c}_i = 0$ (if $\delta\Lambda_i > 0$) or $\tilde{c}_i = 1$ (if $\delta\Lambda_i < 0$). This gives

$$|\hat{b}_i - b_i| \leq |\delta\Lambda_i| \max\left[\hat{\Lambda}_i^{-1}, \Lambda_i^{-1}\right] \tag{75}$$
$$|\hat{b}_i - b_i| \leq 4|\delta\Lambda_i| \tag{76}$$

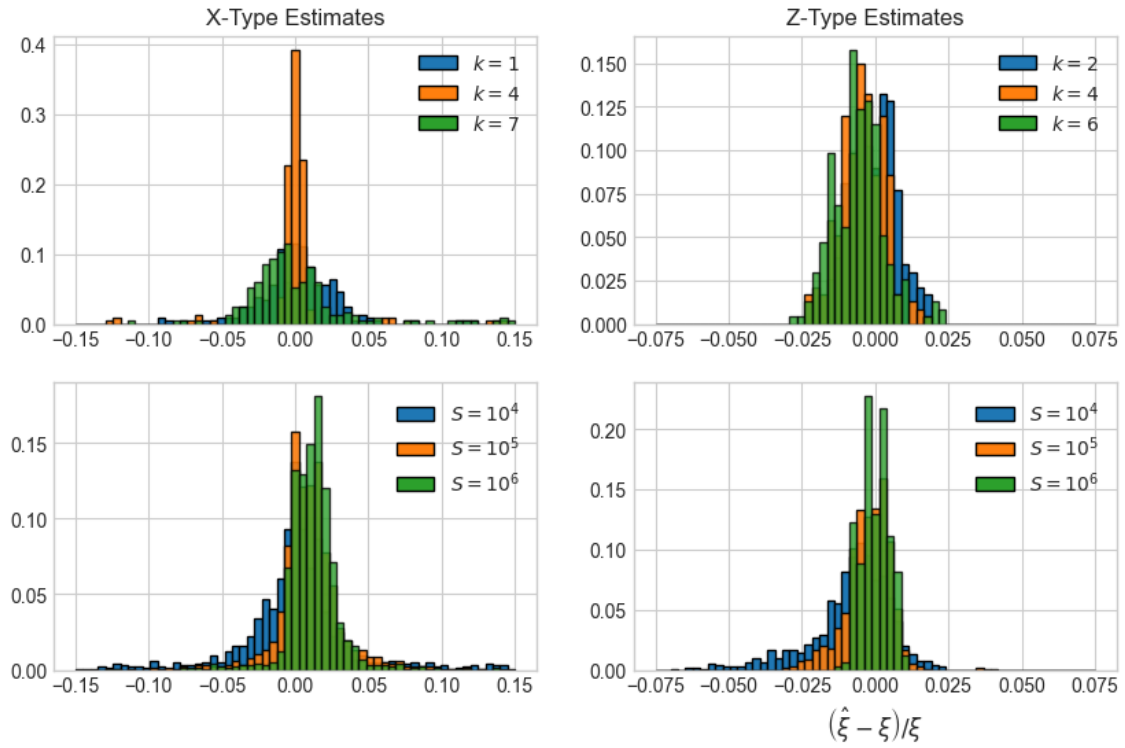


Figure 2: Histogram of relative errors for $2n = 10$ fermionic modes with 1000 circuits for each experiment type. See main text for description.

by our assumption that both $\Lambda_i \geq \frac{1}{2}$ and $\hat{\Lambda}_i \geq \frac{1}{4}$ for all i. Finally, we use this to provide a bound on $\|\hat{\mathbf{b}} - \mathbf{b}\|_\infty$ as

$$\begin{aligned} \|\hat{\mathbf{b}} - \mathbf{b}\|_\infty &\leq 4\|\delta\mathbf{\Lambda}\|_\infty && (77)\\ &= 4\|\mathbf{V}\delta\mathbf{P}\|_\infty && (78)\\ &\leq \max_j \left(4\|\delta\mathbf{P}^{(j)}\|_1\right) && (79)\\ \|\hat{\mathbf{b}} - \mathbf{b}\|_\infty &\leq 4\varepsilon. && (80) \end{aligned}$$

From the second to the third line above, we replaced the maximum over indices i in $\|\cdot\|_\infty$ with a maximization first over each experiment and then within each experiment. We then applied the norm inequality $\|\cdot\|_\infty \leq \|\cdot\|_{1\to\infty}\|\cdot\|_1$ experiment by experiment and used the fact that the maximum absolute element of $\mathbf{V}$ is 1 in each case from eq. (53). Combining eq. (80) with eq. (70) gives the result. □

4 Numerics

Finally, we demonstrate the efficacy of our protocol numerically in fig. 2, where we plot a histogram of relative gate-eigenvalue errors for $n = 5$ qubits ($2n = 10$ fermionic modes) and $m = 1000$ circuits for each experiment type. Our noise model is given by a random 2-qubit Pauli channel following each gate we apply, where the Pauli channel probabilities are chosen uniformly from the interval $[10^{-2} - 10^{-3}, 10^{-2} + 10^{-3}]$. Columns are partitioned according to x- and z-type estimates, and rows are partitioned according to whether the different histograms correspond to different Majorana weight-k or different numbers S of

samples from the Born-rule probability distribution (we assume that we can apply FLO-twirled circuits exactly). We also partition Z-rotation angles into 46 bins of width $\Delta\theta = \frac{\pi}{23}$. We observe that the protocol is able to routinely achieve less than 5% relative error for most of the eigenvalues estimated, and that this steadily improves with the number of shots.

5 Conclusion

We have shown how to learn the matchgate-averaged noise parameters on many matchgates simultaneously via FACES, our fermionic adaptation of the ACES protocol. We envision that this will enable experiments to build a complete characterization of averaged noise parameters for universal gate sets, though several open problems remain. Foremost is the technical challenge of performing matchgate twirling. Unlike Pauli twirling, it seems unlikely that matchgate twirling can be achieved in constant depth. However, there have recently been promising results to show that low-depth circuits can nevertheless approximate Haar random unitaries [25, 26]. If a corresponding result holds for Haar random matchgates, then this issue can be mitigated, if not entirely avoided. Alternatively, we might consider only reproducing the Haar ensemble of FLO circuits up to the second moment, which is sufficient for exactly twirling the channel. It is known that this can be achieved using only gates in the intersection of the matchgate and Clifford groups [27, 28]. Finally, we might employ restricted circuit structure to address this challenge in a similar strategy to interleaved randomized benchmarking (IRB) [29], which has historically enjoyed a great deal of success in practice. In a setting where the circuit structure is limited to the FLO analog of IRB, for example, we expect that FACES will exhibit a "repeated convolution structure" similar to that of matchgate benchmarking [18], since we are considering Haar averages over the group of FLO unitaries.

Another open question involves the notion of a *gauge* freedom in the learned parameters. Formally, such a freedom exists when the design matrix has a nontrivial kernel, whose elements correspond to unlearnable functions of gate eigenvalues. Recent works [11, 12] have shown that these functions correspond to the cut space of a particular graph in the Clifford setting. It is thus germane to ask whether a corresponding characterization is possible in the FLO setting.

It was also shown in [12] that, in the Clifford setting, one can construct design matrices that enable learning all learnable Pauli eigenvalues to *relative* precision. We expect similar results to hold in the setting of FACES, but we leave this open.

Ultimately, we would like to utilize these protocols to not only learn circuit-level noise, but to also mitigate it. With this in mind, an important feature of our protocol is that the design matrix is necessarily block diagonal. In the aforementioned question of gauge parameters, it is already clear that each unlearnable function thus depends only on gate eigenvalues from the same Majorana-degree sector. Additionally, noisy expectation values of fixed Majorana-degree observables are necessarily linear functions of the ideal expectation values, even in the setting where these observables are not Majorana monomials (e.g. measuring the expected energy for a Hamiltonian of fixed Majorana degree). This contrasts with the setting of Pauli observables under Clifford-circuit evolution, and is reminiscent of training methods for error mitigation [30, 31]. We leave a careful consideration of this problem and other applications of our results to quantum simulation for future work.

Acknowledgments

We are grateful to Ashley Montanaro and Ryan O'Donnell for informing us that $\mathbf{M}^{(\ell)}$ was known in the literature as a Kravchuk matrix. Matchgate circuit simulations were performed using the GitHub repository [32].

A Pseudocode Algorithms

In this section we provide pseudocode implementations for the steps of our protocol described in fig. 1. Algorithm 1 is the main function, which learns the gate eigenvalues $\boldsymbol{\xi}$ describing an error model $\mathcal{M}$ for a collection of circuits $\mathcal{C}$ in a given number of measurements per circuit S. First, the design matrix $\mathbf{A}$ is computed using algorithm 2. The rows of this matrix are indexed by circuits in $\mathcal{C}$, and the columns are indexed by noise parameters, such that A_{jk} corresponds to the number of times the noise parameter k appears in circuit j. In algorithm 3, a probability distribution of outcomes is generated for each circuit according to its type (either "x" or "z"). These distributions are then input to algorithm 4, where they are used to calculate the circuit eigenvalues by eq. (48) or eq.s (49) and (50). Finally, this is used to produce the gate eigenvalues $\boldsymbol{\xi}$ using a similar linear inversion procedure to the ACES protocol. In this procedure, $\mathbf{A}$ is truncated to only include its elements which lie above some given cutoff δ. If $\mathbf{A}$ is still full rank, we solve the linear equation $\mathbf{Ax} = \mathbf{b}$ relating the logarithms of the circuit, $\mathbf{b}$, and gate eigenvalues $\mathbf{x}$. Finally, we recover the gate eigenvalues themselves by exponentiating the components of $\mathbf{x}$.

Algorithm 1: FACES$(\mathcal{M}, S)$

Given a FACES model $\mathcal{M} \coloneqq (\mathcal{G}, \boldsymbol{\xi}, \mathcal{C})$ and a number of samples S per circuit, output estimates $\hat{\boldsymbol{\xi}}$ of all the gate eigenvalues $\boldsymbol{\xi}$.

Input: $\mathcal{M}, S$

Compute the design matrix $\mathbf{A}$
$\mathbf{A} \leftarrow$ DESIGN$(\mathcal{M})$
$\mathbf{\Lambda} \leftarrow [\ ]$ ▷ Initialize an empty list
# estimate all circuit eigenvalues
for $\mathcal{U}^{(j)} \in \mathcal{C}$ **do**
    $\boldsymbol{p}^{(j)} \leftarrow$ SAMPLECIRCUIT$(\mathcal{U}^{(j)}, S)$
    $\mathbf{\Lambda}^{(j)} \leftarrow$ CIRCUITEIGS$(\boldsymbol{p}^{(j)})$
    $\mathbf{\Lambda} \leftarrow$ APPEND$(\mathbf{\Lambda}, \mathbf{\Lambda}^{(j)})$
end for
# Estimate the gate eigenvalues
$\hat{\boldsymbol{\xi}} \leftarrow$ GATEEIGS$(\mathbf{\Lambda}, \mathbf{A})$

Output: $\hat{\boldsymbol{\xi}}$

Algorithm 2: DESIGN($\mathcal{M}$)
Compute the design matrix $\mathbf{A}$ for the FACES model $\mathcal{M} \coloneqq (\mathcal{G}, \boldsymbol{\xi}, \mathcal{C})$.

Input: $\mathcal{M}$

$J \leftarrow |\mathcal{C}|$
$K \leftarrow |\mathcal{G}|$
$\mathbf{A} \leftarrow \mathbf{0}_{J \times K}$ ▷ Initialize $\mathbf{A}$ as an all-zeros matrix.
# Compute the design matrix $\mathbf{A}$
for $\mathcal{U}^{(j)} \coloneqq \prod_{m=1}^{L(j)} \mathcal{U}_m^{(j)} \in \mathcal{C}$ **do**
    for $m \in [L(j)]$ **do**
        for $\ell \in [K]$ **do**
            if $\mathcal{U}_m^{(j)} = G_\ell$ **then**
                $A_{j\ell}$++
            end if
        end for
    end for
end for # Check if A has full column-rank.
if RANK(A) $< K$ **then**
    THROW("$\mathbf{A}$ is not full rank.
        $\mathcal{M}$ is not identifiable from $\mathcal{C}$.")
end if

Output: $\mathbf{A}$

Algorithm 3: SAMPLECIRCUIT($\mathcal{U}, S$)
Draw N iid samples from the twirled ensemble for circuit $\mathcal{U}$ and return an estimate $\boldsymbol{p}$ of the measurement outcome probability distribution.

Input: $\mathcal{U}, S$

Set $\boldsymbol{p}$ to all zeros (for circuit type x or z).
$\boldsymbol{p} \leftarrow 0$
for $n \in [N]$ **do**
    $\mathcal{U}' \leftarrow$ TWIRL($\mathcal{U}$) ▷ twirled version of $\mathcal{U}$
    # Run experiment, record single outcome.
    $w \leftarrow$ RUNEXPERIMENT($\mathcal{U}'$)
    $\boldsymbol{p}[w]$++ ▷ Increment count for outcome w
end for

Output: $\boldsymbol{p}/S$

Algorithm 4: CircuitEigs($\boldsymbol{p}$)
Given a probability distribution $\boldsymbol{p}$, compute the associated circuit eigenvalues $\boldsymbol{\Lambda}$ depending on the type of $\boldsymbol{p}$.

Input: $\boldsymbol{p}$

# $\mathbf{M}$ is calculated using theorem 4.
if $\textsc{type}(\boldsymbol{p}) = z$ **then**
# Apply eq. (48)
    $\Lambda_{2k} \leftarrow \sum_{\ell=0}^{n} \binom{n}{\ell}^{-1} M_{k\ell}^{(n)} p_\ell$
else ▷ else $\textsc{type}(\boldsymbol{p}) = x$
Apply eq. (49) and eq. (50)
    $\Lambda_{2k} \leftarrow \sum_{\ell=0}^{n-1} \binom{n-1}{\ell}^{-1} M_{k\ell}^{(n-1)} (p_{+,\ell} + p_{-,\ell})$
    $\Lambda_{2k+1} \leftarrow \sum_{\ell=0}^{n-1} \binom{n-1}{\ell}^{-1} M_{k\ell}^{(n-1)} (p_{+,\ell} - p_{-,\ell})$
end if

Output: $\boldsymbol{\Lambda}$

Algorithm 5: Twirl($\mathcal{U} \coloneqq \prod_{m=1}^{L} \mathcal{U}_m$)

Input: $\mathcal{U}$

$\mathcal{U}' \leftarrow \mathcal{I}$
for $k \in [L]$ **do**
    $\mathcal{V} \leftarrow \textsc{sample}(\mathbb{F}(2n))$
    $\mathcal{U}' \leftarrow \left(\mathcal{U}_k \circ \mathcal{V} \circ \mathcal{U}_k^\dagger\right) \circ \left(\mathcal{U}_k \circ \mathcal{V}^\dagger\right) \circ \mathcal{U}'$
end for

Output: $\mathcal{U}'$

Algorithm 6: GateEigs$(\mathbf{A}, \hat{\mathbf{\Lambda}}, \delta)$

From a design matrix $\mathbf{A}$, a collection of circuit eigenvalue estimates $\hat{\mathbf{\Lambda}}$, and a cutoff parameter δ, return estimates $\hat{\boldsymbol{\xi}}$ of the gate eigenvalues $\boldsymbol{\xi}$.

Input: $\mathbf{A}, \hat{\mathbf{\Lambda}}, \delta$

$J, K \leftarrow \text{SIZE}(\mathbf{A})$

\# Delete $\hat{\mathbf{\Lambda}}$ and associated row of $\mathbf{A}$ if below the cutoff.

$\hat{\mathbf{\Lambda}} \leftarrow \hat{\mathbf{\Lambda}}[\hat{\mathbf{\Lambda}} > \delta]$

$\mathbf{A} \leftarrow \mathbf{A}[\hat{\mathbf{\Lambda}} > \delta, :]$

\# Validate that $\mathbf{A}$ is still full column rank.

if $\text{RANK}(A) < K$ **then**

    THROW("$\mathbf{A}$ is not full rank with cutoff δ.")

end if

$\hat{\mathbf{b}} \leftarrow -\log(\hat{\mathbf{\Lambda}})$

$\hat{\mathbf{x}} \leftarrow \text{SOLVE}(\mathbf{A}, \hat{\mathbf{b}})$ ▷ Solve $\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{b}}$

$\hat{\boldsymbol{\xi}} \leftarrow \exp(-\hat{\mathbf{x}})$

Output: $\hat{\boldsymbol{\xi}}$

B Proofs

B.1 Proof of theorem 4

Lemma 4 (Restatement)**.** *We have the alternate expressions for* $M_{jk}^{(\ell)}$

$$M_{jk}^{(\ell)} = \text{coeff}_k((1-u)^j(1+u)^{\ell-j}) \tag{81}$$

$$= \sum_{\boldsymbol{\alpha}|k,\ell} (-1)^{|\boldsymbol{\alpha}\cap[j]|} \tag{82}$$

$$M_{jk}^{(\ell)} = \sum_{m=\max(0,k+j-\ell)}^{\min(j,k)} \binom{\ell-j}{k-m}\binom{j}{m}(-1)^m . \tag{83}$$

Additionally, $\mathbf{M}^{(\ell)}$ *is proportional to an involution*

$$(\mathbf{M}^{(\ell)})^2 = 2^\ell \mathbf{I}. \tag{84}$$

Proof. eq. (81) immediately follows from the definition eq. (9) and Vieta's formulas.

Now consider eq. (82), and let $\boldsymbol{b} \coloneqq -\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}$ from the definition in eq. (9). We have

$$e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}) = \sum_{\boldsymbol{\alpha}|k,\ell} \prod_{g\in\boldsymbol{\alpha}} b_g \tag{85}$$

$$= \sum_{\boldsymbol{\alpha}|k,\ell} \left(\prod_{g\in\boldsymbol{\alpha}\cap[j]} b_g\right)\left(\prod_{g\in\boldsymbol{\alpha}\setminus[j]} b_g\right) \tag{86}$$

$$e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}) = \sum_{\boldsymbol{\alpha}|k,\ell} (-1)^{|\boldsymbol{\alpha}\cap[j]|}, \tag{87}$$

yielding the result eq. (82).

To show eq. (83), we expand the polynomial in the right side of eq. (81) by the binomial theorem

$$(1+u)^{\ell-j}(1-u)^j = \left[\sum_{m=0}^{\ell-j}\binom{\ell-j}{m}u^m\right]\left[\sum_{m=0}^{j}\binom{j}{m}(-u)^m\right] \tag{88}$$

$$(1+u)^{\ell-j}(1-u)^j = \sum_{k=0}^{\ell}\left[\sum_{m=\max(0,k+j-\ell)}^{\min(j,k)}\binom{\ell-j}{k-m}\binom{j}{m}(-1)^m\right]u^k \tag{89}$$

The limits of the sum are such that the binomial coefficients are well-defined (this is redundant since our convention is such that the binomial coefficients are zero if the arguments are outside of the bounds). The lower limit is such that $m \geq 0$ to satisfy the lower bound of $\binom{j}{m}$ and $k-m \leq \ell-j$ to satisfy the upper bound of $\binom{\ell-j}{k-m}$. The upper bound is such that $m \leq j$ to satisfy the upper bound of $\binom{j}{m}$ and $m \leq k$ to satisfy the lower bound of $\binom{\ell-j}{k-m}$. Equation (89), together with eq. (81), gives the result.

To show eq. (84), let $v_j(u) \equiv u^j$ for $j \in \{0\} \cup [\ell]$. We have, again by eq. (9) and Vieta's formulas,

$$\left[\mathbf{M}^{(\ell)} \cdot \boldsymbol{v}(u)\right]_j = (1+u)^{\ell-j}(1-u)^j \tag{90}$$

$$\mathbf{M}^{(\ell)} \cdot \boldsymbol{v}(u) = (1+u)^\ell \boldsymbol{v}\left(\frac{1-u}{1+u}\right) \tag{91}$$

This gives

$$(\mathbf{M}^{(\ell)})^2 \cdot \boldsymbol{v}(u) = (1+u)^\ell \mathbf{M}^{(\ell)} \cdot \boldsymbol{v}\left(\frac{1-u}{1+u}\right) \tag{92}$$

$$= (1+u)^\ell \left[1+\left(\frac{1-u}{1+u}\right)\right]^\ell \boldsymbol{v}\left[\frac{1-\left(\frac{1-u}{1+u}\right)}{1+\left(\frac{1-u}{1+u}\right)}\right] \tag{93}$$

$$= \left[(1+u)+(1-u)\right]^\ell \boldsymbol{v}\left[\frac{(1+u)-(1-u)}{(1+u)+(1-u)}\right] \tag{94}$$

$$(\mathbf{M}^{(\ell)})^2 \cdot \boldsymbol{v}(u) = 2^\ell \boldsymbol{v}(u). \tag{95}$$

The set of vectors $\{\boldsymbol{v}(u)\}_{u\in\mathbb{C}}$ form a spanning set for $\mathbb{C}^{\ell+1}$, so we must therefore have

$$(\mathbf{M}^{(\ell)})^2 = 2^\ell \mathbf{I}. \tag{96}$$

□

B.2 Proof of theorem 6

Lemma 6 (Restatement)**.** *The Kravchuk matrix and antipode involution commute, and their product has components*

$$(\mathbf{s}^{(\ell)} \cdot \mathbf{M}^{(\ell)})_{jk} = (\mathbf{M}^{(\ell)} \cdot \mathbf{s}^{(\ell)})_{jk} = (-1)^{jk} e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}). \tag{97}$$

Proof. We show that both matrix orderings have the components given by the expression on the right-hand side. First, note

$$(\mathbf{s}^{(\ell)} \cdot \mathbf{M}^{(\ell)})_{jk} = \begin{cases} \mathrm{M}^{(\ell)}_{j,k} & j \text{ even} \\ \mathrm{M}^{(\ell)}_{\ell-j,k} & j \text{ odd} \end{cases}. \tag{98}$$

Since exchanging rows j and $\ell - j$ in $\mathbf{M}^{(\ell)}$ is equivalent to multiplying the argument of the symmetric polynomial in eq. (9) by a factor of -1, we have

$$(\mathbf{s}^{(\ell)} \cdot \mathbf{M}^{(\ell)})_{jk} = e_k[(-1)^j(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j})] \tag{99}$$

$$(\mathbf{s}^{(\ell)} \cdot \mathbf{M}^{(\ell)})_{jk} = (-1)^{jk} e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}) \tag{100}$$

where the second line follows from the identity $e_k(-\boldsymbol{v}) = (-1)^k e_k(\boldsymbol{v})$. To show the result for the other ordering, we utilize the *reciprocal polynomial* of $p(x)$, given by $p^*(x) = x^d p(1/x)$, where d is the degree of p. In particular, the degree-k coefficient of p is the degree-$(d-k)$ coefficient of p^*. For any k, we have by theorem 4

$$\mathrm{M}^{(\ell)}_{j,\ell-k} = \mathrm{coeff}_{\ell-k}((1-u)^j(1+u)^{\ell-j}) \tag{101}$$

$$= \mathrm{coeff}_k(u^\ell(1-\frac{1}{u})^j(1+\frac{1}{u})^{\ell-j}) \tag{102}$$

$$= \mathrm{coeff}_k((u-1)^j(u+1)^{\ell-j}) \tag{103}$$

$$\mathrm{M}^{(\ell)}_{j,\ell-k} = (-1)^j \mathrm{M}^{(\ell)}_{j,k} \tag{104}$$

Since $\mathbf{s}^{(\ell)}$ only exchanges columns k and $\ell - k$ in $\mathbf{M}^{(\ell)}$ if k is odd, we have

$$(\mathbf{M}^{(\ell)} \cdot \mathbf{s}^{(\ell)})_{jk} = (-1)^{jk} \mathrm{M}_{jk} = (-1)^{jk} e_k(-\mathbf{1}_j \oplus \mathbf{1}_{\ell-j}) \tag{105}$$

Together, eq. (100) and eq. (105) complete the proof. □

B.3 Proof of theorem 7

Lemma 7 (Restatement)**.** *Consider a Pauli channel $\mathcal{E}^{\mathbb{P}(n)}$ with Pauli error probabilities $p_{\mathbf{x}}$. Let $\mathbf{x}(\boldsymbol{\alpha})$ denote the Pauli string associated to the subset of Majorana modes $\boldsymbol{\alpha}$ by the Jordan-Wigner transformation. Then $\mathcal{E}^{\mathbb{F}(2n)} = (\mathcal{E}^{\mathbb{P}(n)})^{\mathbb{F}(2n)}$, and the fermionic error probabilities q_k of $\mathcal{E}^{\mathbb{F}(2n)}$ are given by*

$$q_k = \sum_{\boldsymbol{\alpha}|k} p_{\mathbf{x}(\boldsymbol{\alpha})} \,. \tag{106}$$

Proof. The first statement follows from the fact that every Pauli operator is also a FLO unitary by the Jordan-Wigner transformation. Namely, we have, for all $\boldsymbol{\alpha} \subseteq [2n]$ with $|\boldsymbol{\alpha}| = L$,

$$\gamma_{\boldsymbol{\alpha}} = \begin{cases} \prod_{j=1}^{L/2} \left(\gamma_{\alpha_{2j-1}} \gamma_{\alpha_{2j}}\right) & L \text{ even} \\ \gamma_1 \left(\gamma_1 \gamma_{\alpha_1}\right) \prod_{j=1}^{(L-1)/2} \left(\gamma_{\alpha_{2j}} \gamma_{\alpha_{2j+1}}\right) & L \text{ odd} \end{cases} . \tag{107}$$

Each factor in parentheses is a FLO transformation, since $\gamma_j\gamma_k = \exp\left(\frac{\pi}{2}\gamma_j\gamma_k\right)$, and $\gamma_1 \in \mathbb{F}(2n)$ by the definition below eq. (3). Thus, we can perform the FLO twirl on $\mathcal{E}$ by first performing a Pauli twirl, mapping $\mathcal{E} \to \mathcal{E}^{\mathbb{P}(n)}$, and then twirling over the full group of FLO unitary transformation, mapping $\mathcal{E}^{\mathbb{P}(n)} \to (\mathcal{E}^{\mathbb{P}(n)})^{\mathbb{F}(2n)} = \mathcal{E}^{\mathbb{F}(2n)}$.

By applying the Jordan-Wigner transform to eq. (22), together with the first statement of the lemma, we have

$$\mathcal{E}^{\mathbb{F}(2n)}(X) = \sum_{\boldsymbol{\alpha} \subseteq [2n]} p_{\mathbf{x}(\boldsymbol{\alpha})} \int_{\mathbb{F}(2n)} dU \, \mathcal{U}\left(\gamma_{\boldsymbol{\alpha}}\right) X \mathcal{U}(\gamma_{\boldsymbol{\alpha}}^\dagger) \tag{108}$$

$$\mathcal{E}^{\mathbb{F}(2n)}(X) = \sum_{\substack{\boldsymbol{\alpha} \subseteq [2n] \\ \boldsymbol{\beta}, \boldsymbol{\eta} || \boldsymbol{\alpha}|, 2n}} p_{\mathbf{x}(\boldsymbol{\alpha})} \left[\int_{O(2n)} d\mathbf{R} \det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right) \det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\eta}}\right)\right] \left(\gamma_{\boldsymbol{\beta}} X \gamma_{\boldsymbol{\eta}}^\dagger\right) \tag{109}$$

Since $\mathcal{E}^{\mathbb{F}(2n)}$ is invariant under Pauli twirling (again by the argument that every Pauli is a matchgate), we must have that only the terms for which $\boldsymbol{\beta} = \boldsymbol{\eta}$ are nonvanishing. This gives

$$\mathcal{E}^{\mathbb{F}(2n)}(X) = \sum_{\substack{\boldsymbol{\alpha}\subseteq[2n] \\ \boldsymbol{\beta}||\boldsymbol{\alpha}|,2n}} p_{\mathbf{x}(\boldsymbol{\alpha})} \left[\int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right)|^2\right] \left(\gamma_{\boldsymbol{\beta}} \odot \gamma_{\boldsymbol{\beta}}^{\dagger}\right) \tag{110}$$

To evaluate this integral, note that it is invariant upon making the change-of-variables $\mathbf{R} \to \mathbf{S}^T\mathbf{R}$, where $\mathbf{S}$ is a permutation matrix which maps $\boldsymbol{\alpha}$ to some subset $\boldsymbol{\eta}$ with $|\boldsymbol{\eta}| = |\boldsymbol{\alpha}|$, as

$$\int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right)|^2 = \int_{O(2n)} d\mathbf{R} \; |\det\left[\left(\mathbf{S}^{\mathrm{T}}\mathbf{R}\right)_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right]|^2 \tag{111}$$

$$\int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right)|^2 = \int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\eta}\boldsymbol{\beta}}\right)|^2 \tag{112}$$

In this step, we used the Cauchy-Binet formula and the fact that $\det\left[\left(\mathbf{S}^{\mathrm{T}}\right)_{\boldsymbol{\alpha}\boldsymbol{\xi}}\right] = \pm\delta_{\boldsymbol{\xi}\boldsymbol{\eta}}$. Summing over all such $\boldsymbol{\eta}$ gives

$$\int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right)|^2 = \binom{2n}{|\boldsymbol{\alpha}|}^{-1} \sum_{\boldsymbol{\eta}||\boldsymbol{\alpha}|} \int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\eta}\boldsymbol{\beta}}\right)|^2 \tag{113}$$

$$= \binom{2n}{|\boldsymbol{\alpha}|}^{-1} \int_{O(2n)} d\mathbf{R} \; \det\left[\left(\mathbf{R}^{\mathrm{T}}\mathbf{R}\right)_{\boldsymbol{\beta}\boldsymbol{\beta}}\right] \tag{114}$$

$$\int_{O(2n)} d\mathbf{R} \; |\det\left(\mathbf{R}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\right)|^2 = \binom{2n}{|\boldsymbol{\alpha}|}^{-1} \tag{115}$$

From the first to the second equality, we used the fact that $\det\left(\mathbf{R}_{\boldsymbol{\eta}\boldsymbol{\beta}}\right) = \det\left(\mathbf{R}^{\mathrm{T}}_{\boldsymbol{\beta}\boldsymbol{\eta}}\right)$ in one of the factors and used the Cauchy-Binet formula again. Finally, we evaluated the determinant to 1 and used the normalization of the Haar integral. This gives

$$\mathcal{E}^{\mathbb{F}(2n)} = \sum_{\substack{\boldsymbol{\alpha}\subseteq[2n] \\ \boldsymbol{\beta}||\boldsymbol{\alpha}|}} \binom{2n}{|\boldsymbol{\alpha}|}^{-1} p_{\mathbf{x}(\boldsymbol{\alpha})} \left(\gamma_{\boldsymbol{\beta}} \odot \gamma_{\boldsymbol{\beta}}^{\dagger}\right) \tag{116}$$

$$\mathcal{E}^{\mathbb{F}(2n)} = \sum_{k=0}^{2n} \binom{2n}{k}^{-1} \left(\sum_{\boldsymbol{\alpha}|k} p_{\mathbf{x}(\boldsymbol{\alpha})}\right) \left(\sum_{\boldsymbol{\beta}|k} \gamma_{\boldsymbol{\beta}} \odot \gamma_{\boldsymbol{\beta}}^{\dagger}\right) \tag{117}$$

Identifying $q_k = \sum_{\boldsymbol{\alpha}|k} p_{\mathbf{x}(\boldsymbol{\alpha})}$ gives the second statement of the lemma. □

B.4 Proof of theorem 11

Lemma 11 (Restatement)**.** *For an ensemble of noisy and FLO-twirled FLO circuits* $\mathcal{C}_r^{\mathbb{F}(2n)}$*,*

$$\mathbf{P}_0^{(j)} = 2^{-n}\mathbf{d}^{(n)} \cdot \mathbf{M}^{(n)} \cdot \boldsymbol{\Lambda}_{\mathrm{even}}^{(j)} \tag{118}$$

for $r =$ "z"*. Similarly,*

$$\frac{1}{2}(\mathbf{P}_+^{(j)} + \mathbf{P}_-^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \boldsymbol{\Lambda}_{\mathrm{even}\setminus\{2n\}}^{(j)} \tag{119}$$

$$\frac{1}{2}(\mathbf{P}_+^{(j)} - \mathbf{P}_-^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \boldsymbol{\Lambda}_{\mathrm{odd}}^{(j)} \tag{120}$$

for $r =$ "x"*.*

Proof. We begin with the statement for $\mathcal{C}_z^{\mathbb{F}(2n)}$. We have

$$\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) = \prod_{m=1}^{L(j)} \left(\mathcal{E}_m^{(j)}\right)^{\mathbb{F}(2n)} \mathcal{U}_m^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) \tag{121}$$

$$\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) = 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \prod_{m=1}^{L(j)} \left(\mathcal{E}_m^{(j)}\right)^{\mathbb{M}(n)} \mathcal{U}_m^{(j)}(Z^{\mathbf{x}}) \tag{122}$$

where $Z^{\mathbf{x}} \coloneqq \prod_{j=1}^n Z_j^{x_j}$. Under the Jordan-Wigner transform, we have

$$Z^{\mathbf{x}} = \prod_{j=1}^{n} (-i\gamma_{2j-1}\gamma_{2j})^{x_j} \tag{123}$$

Applying the argument leading to eq. (37) gives

$$\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) = 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \prod_{m=1}^{L(j)} \xi_{m,2|\mathbf{x}|}^{(j)} \mathcal{U}_m^{(j)}(Z^{\mathbf{x}}) \tag{124}$$

$$\widetilde{\mathcal{U}}^{(j)}(\rho_0) = 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \Lambda_{2|\mathbf{x}|}^{(j)} \mathcal{U}^{(j)}(Z^{\mathbf{x}}). \tag{125}$$

Circuits in $\mathcal{C}_z^{\mathbb{F}(2n)}$ are chosen to interact as the identity on net. This gives

$$\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|) = 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \Lambda_{2|\mathbf{x}|}^{(j)} Z^{\mathbf{x}} \tag{126}$$

and

$$\langle\mathbf{w}|\,\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|)\,|\mathbf{w}\rangle = 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \Lambda_{2|\mathbf{x}|}^{(j)} \langle\mathbf{w}|\,Z^{\mathbf{x}}\,|\mathbf{w}\rangle \tag{127}$$

$$= 2^{-n} \sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}} \Lambda_{2|\mathbf{x}|}^{(j)} (-1)^{\mathbf{x}\cdot\mathbf{w}} \tag{128}$$

$$= 2^{-n} \sum_{k=0}^{n} \Lambda_{2k}^{(j)} \sum_{\mathbf{x}|k} (-1)^{\mathbf{x}\cdot\mathbf{w}} \tag{129}$$

$$\langle\mathbf{w}|\,\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|)\,|\mathbf{w}\rangle = 2^{-n} \sum_{k=0}^{n} \Lambda_{2k}^{(j)} M_{|\mathbf{w}|k} \tag{130}$$

$$\tag{131}$$

Summing over all $\mathbf{w}$ with a fixed Hamming weight ℓ gives

$$\sum_{\mathbf{w}|\ell} \langle\mathbf{w}|\,\widetilde{\mathcal{U}}^{(j)}(|\mathbf{0}\rangle\langle\mathbf{0}|)\,|\mathbf{w}\rangle = 2^{-n} \binom{n}{\ell} \sum_{k=0}^{n} M_{\ell k} \Lambda_{2k}^{(j)} \tag{132}$$

which is the desired result.

$$\mathrm{P}_0^{(j)} = 2^{-n} \mathbf{d}^{(n)} \cdot \mathbf{M}^{(n)} \cdot \mathbf{\Lambda}_{\mathrm{even}}^{(j)}. \tag{133}$$

In the $\mathcal{C}_x^{\mathbb{F}(2n)}$-case, instead of eq. (122), we have

$$\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n}) = \prod_{m=1}^{L(j)} \left(\mathcal{E}_m^{(j)}\right)^{\mathbb{F}(2n)} \mathcal{U}_m^{(j)}(|+\rangle\langle+|^{\otimes n}) \tag{134}$$

As shown in [18], there exists a FLO unitary $\mathcal{U}_+$ effecting

$$\mathcal{U}_+(|+\rangle\langle+|^{\otimes n}) = \left(\frac{I+i\gamma_1}{\sqrt{2}}\right)|\mathbf{0}\rangle\langle\mathbf{0}|\left(\frac{I-i\gamma_1}{\sqrt{2}}\right) \tag{135}$$

This gives

$$\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n}) = \prod_{m=1}^{L(j)}\left(\mathcal{E}_m^{(j)}\right)^{\mathbb{F}(2n)}\mathcal{U}_m^{(j)}\left\{\mathcal{U}_+^\dagger\left[\left(\frac{I+i\gamma_1}{\sqrt{2}}\right)|\mathbf{0}\rangle\langle\mathbf{0}|\left(\frac{I-i\gamma_1}{\sqrt{2}}\right)\right]\right\} \tag{136}$$

$$= 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}}\prod_{m=1}^{L(j)}\left(\mathcal{E}_m^{(j)}\right)^{\mathbb{F}(2n)}\mathcal{U}_m^{(j)}\left\{\mathcal{U}_+^\dagger\left[\left(\frac{I+i\gamma_1}{\sqrt{2}}\right)Z^{\mathbf{x}}\left(\frac{I-i\gamma_1}{\sqrt{2}}\right)\right]\right\} \tag{137}$$

$$= 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}}\prod_{m=1}^{L(j)}\left(\mathcal{E}_m^{(j)}\right)^{\mathbb{F}(2n)}\mathcal{U}_m^{(j)}\left\{\mathcal{U}_+^\dagger\left[(i\gamma_1)^{x_1}Z^{\mathbf{x}}\right]\right\} \tag{138}$$

$$= 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}}\prod_{m=1}^{L(j)}\xi_{m,2|\mathbf{x}|-x_1}^{(j)}\mathcal{U}_m^{(j)}\left\{\mathcal{U}_+^\dagger\left[(i\gamma_1)^{x_1}Z^{\mathbf{x}}\right]\right\} \tag{139}$$

$$\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n}) = 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}}\Lambda_{2|\mathbf{x}|-x_1}^{(j)}\prod_{m=1}^{L(j)}\mathcal{U}_m^{(j)}\left\{\mathcal{U}_+^\dagger\left[(i\gamma_1)^{x_1}Z^{\mathbf{x}}\right]\right\}. \tag{140}$$

Since these circuits act as $\mathcal{U}_+$ on-net, this gives

$$\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n}) = 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times n}}\Lambda_{2|\mathbf{x}|-x_1}^{(j)}(i\gamma_1)^{x_1}Z^{\mathbf{x}} \tag{141}$$

Measuring in the Y-basis on the first qubit and the Z-basis on all other qubits, we have

$$\langle+_y\mathbf{y}|\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n})|+_y\mathbf{y}\rangle = 2^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_2^{\times(n-1)}}\left(\Lambda_{2|\mathbf{x}|}^{(j)}+\Lambda_{2|\mathbf{x}|+1}^{(j)}\right)(-1)^{\mathbf{x}\cdot\mathbf{y}} \tag{142}$$

$$= 2^{-n}\sum_{k=0}^{n-1}\left(\Lambda_{2k}^{(j)}+\Lambda_{2k+1}^{(j)}\right)\sum_{\mathbf{x}|k}(-1)^{\mathbf{x}\cdot\mathbf{y}} \tag{143}$$

$$\langle+_y\mathbf{y}|\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n})|+_y\mathbf{y}\rangle = 2^{-n}\sum_{k=0}^{n-1}\left(\Lambda_{2k}^{(j)}+\Lambda_{2k+1}^{(j)}\right)M_{|\mathbf{y}|k}^{(n-1)}. \tag{144}$$

Summing again over all probabilities of a fixed Hamming weight gives

$$\sum_{\mathbf{y}|\ell,n-1}\langle+_y\mathbf{y}|\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n})|+_y\mathbf{y}\rangle = 2^{-n}\sum_{k=0}^{n-1}\left(\Lambda_{2k}^{(j)}+\Lambda_{2k+1}^{(j)}\right)\sum_{\mathbf{y}|\ell,n-1}M_{|\mathbf{y}|k}^{(n-1)} \tag{145}$$

$$\sum_{\mathbf{y}|\ell,n-1}\langle+_y\mathbf{y}|\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n})|+_y\mathbf{y}\rangle = 2^{-n}\binom{n-1}{\ell}\sum_{k=0}^{n-1}\left(\Lambda_{2k}^{(j)}+\Lambda_{2k+1}^{(j)}\right)M_{\ell k}^{(n-1)} \tag{146}$$

The corresponding probability for the "$-$" outcome on the first bit is given by

$$\sum_{\mathbf{y}|\ell,n-1}\langle-_y\mathbf{y}|\widetilde{\mathcal{U}}^{(j)}(|+\rangle\langle+|^{\otimes n})|-_y\mathbf{y}\rangle = 2^{-n}\binom{n-1}{\ell}\sum_{k=0}^{n-1}\left(\Lambda_{2k}^{(j)}-\Lambda_{2k+1}^{(j)}\right)M_{\ell k}^{(n-1)} \tag{147}$$

and this gives the desired relations

$$\frac{1}{2}(\mathrm{P}_{+}^{(j)} + \mathrm{P}_{-}^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \mathbf{\Lambda}_{\mathrm{even}\setminus\{2n\}}^{(j)} \tag{148}$$

$$\frac{1}{2}(\mathrm{P}_{+}^{(j)} - \mathrm{P}_{-}^{(j)}) = 2^{-n}\mathbf{d}^{(n-1)} \cdot \mathbf{M}^{(n-1)} \cdot \mathbf{\Lambda}_{\mathrm{odd}}^{(j)} \tag{149}$$

By estimating these probabilities, we can collectively estimate all of the noise eigenvalues. □

References

[1] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, et al. "Quantum supremacy using a programmable superconducting processor". Nature **574**, 505–510 (2019).

[2] Guido Burkard, Thaddeus D. Ladd, Andrew Pan, John M. Nichol, and Jason R. Petta. "Semiconductor spin qubits". Rev. Mod. Phys. **95**, 025003 (2023).

[3] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. "Quantum computational chemistry". Rev. Mod. Phys. **92**, 015003 (2020).

[4] E. Knill. "Quantum computing with realistically noisy devices". Nature **434**, 39–44 (2005).

[5] Joel J. Wallman and Joseph Emerson. "Noise tailoring for scalable quantum computation via randomized compiling". Phys. Rev. A **94**, 052325 (2016).

[6] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Jonathan Mizrahi, Jonathan D. Sterk, and Peter Maunz. "Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit" (2013). arXiv:1310.4492.

[7] Dohun Kim, D. R. Ward, C. B. Simmons, John King Gamble, Robin Blume-Kohout, Erik Nielsen, D. E. Savage, M. G. Lagally, Mark Friesen, S. N. Coppersmith, and M. A. Eriksson. "Microwave-driven coherent operation of a semiconductor quantum dot charge qubit". Nature Nanotechnology **10**, 243–247 (2015).

[8] Erik Nielsen, John King Gamble, Kenneth Rudinger, Travis Scholten, Kevin Young, and Robin Blume-Kohout. "Gate Set Tomography". Quantum **5**, 557 (2021).

[9] Steven T. Flammia. "Averaged Circuit Eigenvalue Sampling". In François Le Gall and Tomoyuki Morimae, editors, 17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022). Volume 232 of Leibniz International Proceedings in Informatics (LIPIcs), pages 4:1–4:10. Dagstuhl, Germany (2022). Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[10] Richard Jozsa and Akimasa Miyake. "Matchgates and classical simulation of quantum circuits". Proc. R. Soc. A. **464**, 3089–3106 (2008).

[11] Senrui Chen, Yunchao Liu, Matthew Otten, Alireza Seif, Bill Fefferman, and Liang Jiang. "The learnability of Pauli noise". Nat Commun **14**, 52 (2023).

[12] Senrui Chen, Zhihan Zhang, Liang Jiang, and Steven T. Flammia. "Efficient self-consistent learning of gate set pauli noise". PRX Quantum **7**, 010305 (2026).

[13] Evan T. Hockings, Andrew C. Doherty, and Robin Harper. "Scalable Noise Characterization of Syndrome-Extraction Circuits with Averaged Circuit Eigenvalue Sampling". PRX Quantum **6**, 010334 (2025).

[14] Jahan Claes, Eleanor Rieffel, and Zhihui Wang. "Character randomized benchmarking for non-multiplicity-free groups with applications to subspace, leakage, and matchgate randomized benchmarking". PRX Quantum **2**, 010351 (2021).

[15] Guoding Liu, Ziyi Xie, Zitai Xu, and Xiongfeng Ma. "Group twirling and noise tailoring for multiqubit controlled phase gates". Phys. Rev. Res. **6**, 043221 (2024).

[16] Jadwiga Wilkens, Marios Ioannou, Ellen Derbyshire, Jens Eisert, Dominik Hangleiter, Ingo Roth, and Jonas Haferkamp. "Benchmarking bosonic and fermionic dynamics" (2024). arXiv:2408.11105.

[17] Joshua Cudby and Sergii Strelchuk. "Learning Gaussian Operations and the Matchgate Hierarchy" (2024). arXiv:2407.12649.

[18] Jonas Helsen, Sepehr Nezami, Matthew Reagor, and Michael Walter. "Matchgate benchmarking: Scalable benchmarking of a continuous family of many-qubit gates". Quantum **6**, 657 (2022).

[19] Jędrzej Burkat and Sergii Strelchuk. "A Lightweight Protocol for Matchgate Fidelity Estimation" (2024). arXiv:2404.07974.

[20] Christoph Dankert. "Efficient Simulation of Random Quantum States and Operators" (2005). arXiv:quant-ph/0512217.

[21] Zhang Jiang, Kevin J. Sung, Kostyantyn Kechedzhi, Vadim N. Smelyanskiy, and Sergio Boixo. "Quantum algorithms to simulate many-body physics of correlated fermions". Phys. Rev. Appl. **9**, 044036 (2018).

[22] D S França and A K Hashagen. "Approximate randomized benchmarking for finite groups". J. Phys. A: Math. Theor. **51**, 395302 (2018).

[23] Steven T. Flammia and Joel J. Wallman. "Efficient estimation of pauli channels". ACM Transactions on Quantum Computing **1**, 1–32 (2020).

[24] Seth T. Merkel, Jay M. Gambetta, John A. Smolin, Stefano Poletto, Antonio D. Córcoles, Blake R. Johnson, Colm A. Ryan, and Matthias Steffen. "Self-consistent quantum process tomography". Phys. Rev. A **87**, 062119 (2013).

[25] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. "Random unitaries in extremely low depth" (2024). arXiv:2407.07754.

[26] Nicholas LaRacuente and Felix Leditzky. "Approximate unitary k-designs from shallow, low-communication circuits". Commun. Math. Phys. **407** (2026).

[27] Kianna Wan, William J. Huggins, Joonho Lee, and Ryan Babbush. "Matchgate shadows for fermionic quantum simulation". Commun. Math. Phys. **404**, 629–700 (2023).

[28] Valentin Heyraud, Héloise Chomet, and Jules Tilly. "Unified Framework for Matchgate Classical Shadows" (2024). arXiv:2409.03836.

[29] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. "Efficient measurement of quantum gate error by interleaved randomized benchmarking". Phys. Rev. Lett. **109**, 080505 (2012).

[30] Ashley Montanaro and Stasja Stanisic. "Error mitigation by training with fermionic linear optics" (2021). arXiv:2102.02120.

[31] Miha Papič, Manuel G. Algaba, Emiliano Godinez-Ramirez, Inés de Vega, Adrian Auer, Fedor Šimkovic IV, and Alessio Calzona. "Near-Term Fermionic Simulation with Subspace Noise Tailored Quantum Error Mitigation" (2025). arXiv:2503.11785.

[32] Jahan Claes (2020). code: jahanclaes/Hoffman-Decomposition-and-the-Matchgate-Group.git.