Fermion Sampling: A Robust Quantum Computational Advantage Scheme Using Fermionic Linear Optics and Magic Input States

Michał Oszmaniec,[1,*] Ninnat Dangniam,[1] Mauro E.S. Morales,[2] and Zoltán Zimborás[3,4,5,†]

[1] *Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, Warszawa 02-668, Poland*
[2] *Centre for Quantum Computation and Communication Technology, Centre for Quantum Software and Information, University of Technology Sydney, NSW 2007, Australia*
[3] *Wigner Research Centre for Physics, Budapest H-1121, Hungary*
[4] *MTA-BME Lendület Quantum Information Theory Research Group Budapest, Hungary*
[5] *Mathematical Institute, Budapest University of Technology and Economics, Budapest H-1111, Hungary*

(Received 21 January 2021; accepted 25 March 2022; published 9 May 2022)

Fermionic linear optics (FLO) is a restricted model of quantum computation, which in its original form is known to be efficiently classically simulable. We show that, when initialized with suitable input states, FLO circuits can be used to demonstrate quantum computational advantage with strong hardness guarantees. Based on this, we propose a quantum advantage scheme, which is a fermionic analog of boson sampling: fermion sampling with magic input states. We consider in parallel two classes of circuits: particle-number conserving (passive) FLO and active FLO that preserves only fermionic parity. Using low-dimensional continuous symmetry groups that underpin these classes of quantum circuits, we prove anticoncentration and robust average-case hardness of computation of output probabilities probabilities. Taken together, these findings provide hardness guarantees comparable to the paradigm of random circuit sampling and boson sampling, the leading candidates for attaining quantum computational advantage. Our scheme is experimentally feasible. FLO circuits are relevant for quantum chemistry and many-body physics, and have been successfully implemented in superconducting architectures. We also argue that due to the structured nature of FLO circuits, they can be efficiently certified using resources scaling polynomially with the system size, with partial trust in the quantum device.

DOI: 10.1103/PRXQuantum.3.020328

I. INTRODUCTION

Universal fault-tolerant quantum computers are expected to exceed capabilities of classical computers in many applications including optimization problems, simulation of many-body quantum systems, machine learning, and code breaking. However, practical requirements for implementations of quantum algorithms generally require the noise level to be below a certain stringent threshold and an encoding of logical qubits into a large number of physical qubits [1]. Despite the impressive progress made along the road to realize a large-scale fault-tolerant quantum computer as shown in the proof-of-principle demonstrations of error correction [2] and fault tolerance [3], what we have at present and in the near future are noisy intermediate-scale quantum (NISQ) devices [4]: NISQ processors having the order of tens or hundreds of qubits.

The paradigm of quantum computational advantage [7,8] (also known as supremacy) aims to develop schemes showing computational advantage of restricted-purpose quantum machines under minimal theoretical assumptions while minimizing hardware requirements. Importantly, given the current status of complexity theory, a rigorous separation of the power of quantum and classical computers cannot be made without plausible assumptions such as the noncollapse of the polynomial hierarchy (a weaker version of $P \neq NP$). Current quantum advantage schemes are usually based on the problem of sampling, i.e., the task of generating samples of a distribution generated by a given quantum circuit or a specifically tuned device (see, however, Ref. [9] for an alternative proposal involving *relation problems* that challenges classical computers in the playground of shallow circuits). The first candidate for demonstration of quantum computational supremacy was boson

[*] oszmaniec@cft.edu.pl
[†] zimboras.zoltan@wigner.hu

sampling [10] that proposed to sample from photonic networks that were initialized in single-photon photon states of several modes. Subsequent sampling proposals include the instantaneous quantum polynomial (IQP) sampling [11,12], the random circuit sampling (RCS) [13–15], the quantum Fourier sampling [16], and many other schemes [17–21] that usually operate on multiqubit systems that undergo evolutions under restricted gate sets (there exist, however, other proposals that use Gaussian states [22,23] or atomic systems [18,24]). Currently random circuit sampling and boson sampling are considered to be the most promising candidates for demonstration of quantum supremacy, both in terms of experimental feasibility and theoretical hardness guarantees.

RCS is the task of sampling from the output distribution of a randomly selected quantum circuit. RCS has been recently experimentally demonstrated in a system of 53 superconducting qubits arranged in the planar layout, and using random two-qubit-local circuits of depth 20 [25]. Beside its experimental feasibility in current NISQ architectures, this quantum advantage proposal also enjoys strong hardness guarantees based on two technical results available for random quantum circuits: a worst-to-average-case reduction of the hardness of computing the output probabilities [14,15] and anticoncentration [26,27].

Independently of RCS, the original proposal of boson sampling received a lot of interest due to the experimental progress in the field of integrated photonics [28]. Currently, the state-of-the-art experiments involve 14 indistinguishable photons in 20 modes [29], while a recent work [30] reported demonstration of Gaussian boson sampling (i.e., a variant of boson sampling with Gaussian input states) using 50 squeezed states at the input of a 100-mode photonic network. Worst-to-average-case reduction for the exact computation of the probabilities in boson sampling was proven in Ref. [10]. Shortly after completion of this work worst-to-average-case reductions were provided for approximate computation of outcome probabilities in boson sampling [31] and Gaussian boson sampling [32]. The anticoncentration property remains unproven for both schemes.

In this work, we propose a quantum advantage scheme based on a fermionic analog of boson sampling: fermion sampling with magic input states. In our scheme a suitable input state $|\Psi_{\mathrm{in}}\rangle$ in $d = 4N$ fermionic modes is transformed via fermionic linear optical (FLO) transformation V, and is measured using particle-number resolving detectors (see Fig. 1). We consider in parallel two classes of circuits: particle-number conserving (passive) FLO and active FLO that preserves only the fermionic parity [33,34] and is closely related to matchgate circuits introduced by Valiant [35]. Mathematically, these classes of circuits can be understood as fermionic representations of the Lie groups $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$. This observation allows us to prove our main technical results. We first show anticoncentration for probabilities in random FLO circuits of both kinds. Moreover, we prove robust average-case hardness of computation of probabilities. To achieve this we adapt the worst-to-average-case reduction based on Cayley transform [15] to our scenario, when instead of the defining representation of the unitary group one considers higher-dimensional representations of low-dimensional Lie groups. Taken together, these findings give hardness guarantees matching that of the paradigm of RCS and boson sampling. We also argue that, due to the structural properties of FLO gates, one can efficiently certify them with resources scaling polynomially with the system size, assuming partial trust in the quantum device.

We argue that our scheme is feasible to realize experimentally. While experimental realization of linear-optical transformation in systems of *real* fermionic systems is

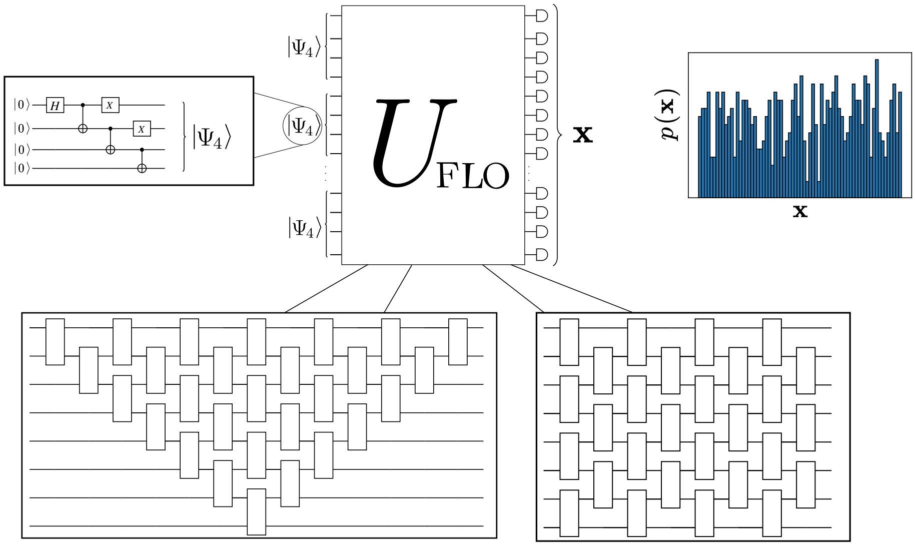


FIG. 1. The setup considered in our work. We run a FLO circuit U_{FLO} (passive or active) with input state $|\Psi_{\mathrm{in}}\rangle = |\Psi_4\rangle^{\otimes N}$ and sample bitstrings $\mathbf{x}$ with the probability distribution $p(\mathbf{x})$ induced by the circuit. Using Jordan-Wigner transformation that encodes fermions in qubits, the state $|\Psi_4\rangle$ can be easily prepared as shown in the inset to the left. The decomposition of the circuits into an elementary gate set can be realized by the fermionic analogs of existing layouts for linear optical networks [5,6] as discussed in Appendix A.

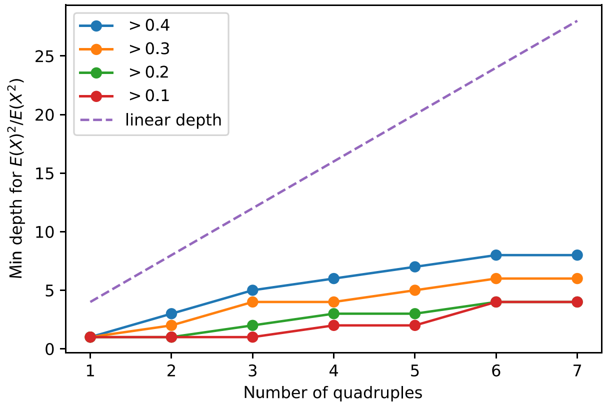


FIG. 2. Plots for the minimum depth required in circuits of a fixed number of quadruples for the ratio $(\mathbb{E}X)^2/\mathbb{E}X^2$, with $X = |\langle \mathbf{x}| V|\Psi_{\text{in}}\rangle|^2$ (V is the random circuit and $|\Psi_{\text{in}}\rangle$ is the quadruple input) to surpass a threshold defined in the legend. For comparison the linear depth in the number of modes is shown. The data suggests that the depths required to obtain the anticoncentration property scale sublinearly.

usually hard because of Coulomb interaction (see, however, Ref. [36]), we make use of the fact that this class of operations is relevant for performing quantum chemistry and many-body simulations on a quantum computer [37–40]. Specifically, after a standard Jordan-Wigner encoding of qubits into fermions, our sampling proposal becomes readily implementable by restricted set of gates and layouts native to superconducting qubit architecture used in simulations of quantum chemistry [41]. A generic FLO transformation can in this way be implemented in depth $\propto N$ in contrast to this the RCS scheme can be implemented in depth $\sqrt{N}$. In the Jordan-Wigner encoding, the magic input states can be prepared using three entangling gates per each and every disjoint block of four qubits and particle-number measurements are realized via standard computational basis measurement (see Fig. 1). While using circuits of linear depth might seem challenging at first sight, this requirement follows solely from our proof techniques that guarantee anticoncentration when sampling from uniform distribution on passive and active FLO circuits. However, we give numerical evidence outcome probabilities corresponding to active FLO circuits anticoncentrate in much smaller depth (see Fig. 2). In fact, it is plausible that active FLO circuits' anticoncentrate in *logarithmic* depth, just like random quantum circuits formed from universal gates, as proved by the recent work by Dalzell *et al.* [42].

A. Significance of results and relation to prior work

1. Relevance of the technical results for hardness of fermion sampling

A first step in establishing hardness of any quantum advantage proposal is showing hardness of sampling up to relative error. Sampling in relative error refers to the task in which a classical computer, given a classical description of the quantum process of interest (e.g., input states, arrangement of gates in the device, etc.), is challenged to efficiently sample from the probability distribution $\{q_{\mathbf{x}}\}$ that for every output $\mathbf{x}$ satisfies $|p_{\mathbf{x}} - q_{\mathbf{x}}| \leq \alpha p_{\mathbf{x}}$, where $\{p_{\mathbf{x}}\}$ is the true probability distribution produced by the device and $\alpha > 0$ is a constant. To establish hardness of sampling up to some relative error, it suffices to show that certain probabilities produced by the device are $\#P$-hard to compute [47], assuming that polynomial hierarchy does not collapse. This hardness of quantum probabilities for specific circuits and outcomes (i.e., in the worst case) was proven long ago for a circuit built from universal gates [48,49]. For nonuniversal models of quantum computation a standard technique for establishing $\#P$-hardness of computation of probabilities is based on showing that a particular nonuniversal model becomes universal when postselection is allowed [10,11, 18,19,50–52]. Relative error approximation is however too strong to be a reasonable notion of approximation from the physical perspective. This is because even a very small amount of experimental noise can render very large relative error.

A more realistic notion of approximate sampling is based on additive error [10,20] in which classical computer is supposed to efficiently produce samples from probability distribution $\{q_{\mathbf{x}}\}$ satisfying $\sum_{\mathbf{x}} |p_{\mathbf{x}} - q_{\mathbf{x}}| \leq \epsilon$, where ϵ is the error parameter. Establishing hardness for additive error approximate sampling is however much more challenging than in the case of relative error. Assuming noncollapse of the polynomial hierarchy, the currently existing techniques [10,12,20] establish this hardness using Stockmayer approximate counting algorithm [53] and by relying on two technical properties of a given quantum advantage proposal: (i) anticoncentration of outcome probabilities, and (ii) $\#P$-hardness of relative error approximate computation of outcome probabilities on average. Anticoncentration refers to property that probability amplitudes $p_{\mathbf{x}}(V)$ are typically not too small, compared to their average value, for random circuits V defining a given quantum advantage proposal. Anticoncentration property has been shown in several schemes [12,17,19,24,27], while for others, including Fourier sampling [16] and boson sampling [10] it remains unproven. On the other hand, average-case $\#P$-hardness of relative error approximate computation of $p_{\mathbf{x}}(V)$ has not been proven to date for the existing quantum advantage proposals. There however exist intermediate results that support it in the form of average-case $\#P$-hardness of *exact* computation of $p_{\mathbf{x}}(V)$ for boson sampling [10], RCS [14], and related schemes [24]. These works adopt the polynomial interpolation technique from Ref. [10] and to prove worst-to-average-case hardness reduction. This reduction has been recently improved by Movassagh [15] for RCS who showed that it is average-case #P-hard to approximate $p_{\mathbf{x}}(V)$ in additive error $\exp[-\Theta(N^{4.5})]$, where N is the number of qubits.

In this work, in order to justify computational hardness of the proposed fermion sampling scheme we prove the following results:

(i) Anticoncentration of probabilities $p_{\mathbf{x}}(V)$ in the output of the scheme for both passive and active FLO circuits initialized in magic states.

(ii) Robust worst-to-average-case hardness reduction for computation of probabilities for passive and active FLO circuits initialized in magic states up to error $\exp[-\Theta(N^6)]$.

Instrumental to our proofs is the fact that active and passive FLO circuits are representations of the low-dimensional (of dimensions scaling polynomially with the number of fermionic modes d) Lie groups $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$, respectively. For the anticoncentration property, we do not use the 2-design property (which is not satisfied for FLO unitaries), but instead prove it relying on specific group-theoretic properties of FLO circuits. For the worst-to-average-case reduction, we follow the state-of-the-art technique by Movassagh [15], which utilizes Cayley path to construct a low-degree rational interpolation between the worst-case and average-case circuits, while generalizing it in two significant directions. First, while the interpolation in Ref. [15] is performed directly using physical circuits, ours is performed at the level of group elements, which are then represented as circuits (see Fig. 3). Secondly, while Ref. [15] applies the interpolation to local one- and two-qubit gates that constitute the circuit, we directly apply it to a global circuit while maintaining the low-degree nature of the rational functions, which is required for the robust reduction.

These results put fermion sampling at the comparable level as RCS [14,15] in terms of state-of-the-art hardness guarantees, surpassing that of boson sampling. The advantage of our scheme compared to RCS is that FLO circuits can be efficiently certified due to their low-dimensional structural properties. The apparent disadvantage is the size of the required circuits—RCS can be implemented in depth $\sqrt{N}$ [13,27], while our scheme requires depth of the structured circuit scaling like N. A more comprehensive comparison of fermionic sampling scheme with other proposals in the literature is given in Table I. We compare the schemes in terms of guarantees for hardness sampling: average-case hardness, anticoncentration, as well as experimental results implementing the schemes or similar results.

2. Comparison with boson sampling

Boson sampling [10], the first quantum advantage proposal based on sampling, relies on the fact that the probability amplitudes of indistinguishable bosons initially prepared in a Fock state and passing through linear-optical network, can be expressed via matrix permanents. Computation of permanent is know to be #P-hard in the worst case [54]. In contrast, the analogous amplitudes for fermions are given by the determinant, which can be computed efficiently. Physically, this difference in complexity can be attributed to the fact that bosonic Fock states are *non-Gaussian* bosonic states, while their fermionic counterparts are in fact fermionic Gaussian states [55]. Thus, to make a closer analogy with boson sampling, we define our fermion sampling using *non-Gaussian* input states $|\Psi_{\mathrm{in}}\rangle = |\Psi_4\rangle^{\otimes N}$, where $|\Psi_4\rangle = 1/\sqrt{2}(|0011\rangle + |1100\rangle)$. This state can be prepared easily on a quantum computer but at the same time can be expressed as an exponential sum

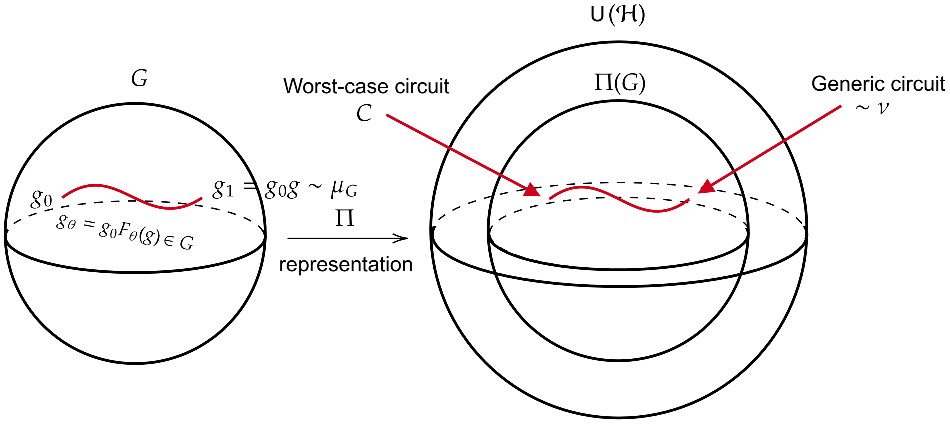
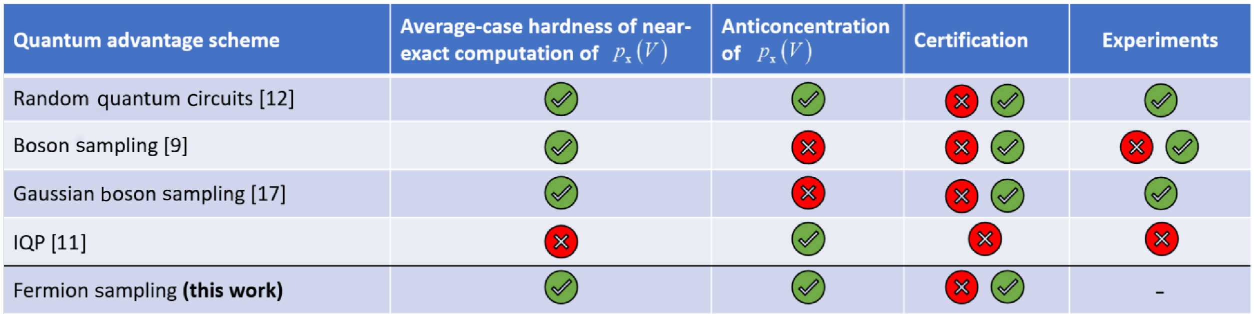


FIG. 3. Π is a group representation from G to [a subgroup $\Pi(G)$ of] the group $\mathrm{U}(\mathcal{H})$ of all quantum circuits on Hilbert space $\mathcal{H}$ (typically of exponential dimension). The Cayley path $g_\theta = g_o F_\theta(g) \in G$ gives a rational interpolation between a fixed element g_0 and $g_1 = g_0 g$. This gives rise to a rational interpolation between circuits $C := \Pi(g_0)$ and $\Pi(g_0)\Pi(g) = \Pi(g_0 g)$. To carry out worst-to-average-case reduction we consider g_0 to be group element corresponding to the worst-case circuit C while g is chosen to be a *generic* element of the Lie group G.

TABLE I. Summarized information for some sampling schemes proposed for quantum supremacy. The first column lists the different sampling schemes: random circuit sampling (RCS) [13,14,25,26,43,44], boson sampling (BS) [10,28,29,31], Gaussian boson sampling (GBS) [22,30,45], instantaneus quantum polynomial (IQP) [12,46], and fermion sampling. The second and third columns indicate if worst-case and average-case hardness for computing the output probabilities has been proven in the approximate case. The third column indicates if the circuits involved in the scheme fulfill the anticoncentration property, the fourth column shows if the scheme has some certification procedure for the circuits involved and the final column indicates if there are experiments that realize the sampling scheme. The boson sampling entry on experiments represents the fact that the current experimental setups are proof-of-principle experiments.



Quantum advantage scheme	Average-case hardness of near-exact computation of $p_x(V)$	Anticoncentration of $p_x(V)$	Certification	Experiments
Random quantum circuits [12]	✓	✓	✗ ✓	✓
Boson sampling [9]	✓	✗	✗ ✓	✗ ✓
Gaussian boson sampling [17]	✓	✗	✗ ✓	✓
IQP [11]	✗	✓	✗	✗
Fermion sampling **(this work)**	✓	✓	✗ ✓	-

of orthogonal Fock states. This is sufficient to guarantee hardness of the corresponding probability amplitudes. It was shown by Ivanov and Gurvits [56,57] that if $|\Psi_{\rm in}\rangle$ is transformed via particle-number preserving (passive) FLO transformation, the probability amplitudes are related to *mixed discriminants* of matrices, which is known to be #*P*-hard, because they can be efficiently reduced to permanent. In the context of active FLO transformations, auxiliary states $|\Psi_4\rangle$ are known to promote this class of transformations to universality [58] (see also Ref. [59]), which can be used to show #*P*-hardness of probabilities arising from active FLO circuits initialized with such non-Gaussian states. We conclude the comparison with boson sampling by clarifying the role of the measurements used. Our proposal uses fermionic particle-number measurements, which are themselves fermionic Gaussian. This differentiates fermion sampling from boson sampling schemes. This includes Gaussian boson sampling [22] in which bosonic squeezed states (that are bosonic Gaussian) are transformed using linear optics, and finally measured using (non-Gaussian) particle-number detectors. In that proposal non-Gaussian character of the particle-number measurement is crucial for hardness [60].

3. Comparison to existing benchmarking protocols

A convincing demonstration of quantum supremacy requires a means to build confidence that the output $q(\mathbf{x})$ of the quantum device is close to the ideal distribution $p(\mathbf{x})$. Such verification can be done with differing levels of efficiency depending on the level of trust in the functioning of the device. Of the highest standard (requiring minimal assumptions) of such verification is to certify whether $q(\mathbf{x}) = p(\mathbf{x})$ or $\sum_{\mathbf{x}} |p_{\mathbf{x}} - q_{\mathbf{x}}| > \epsilon$ for some small $\epsilon > 0$—that is, we are ruling out *all* adversarial distributions that are ϵ away from $p(\mathbf{x})$—using only the classical output statistics of the sampling device. Building on a classical result on identity testing of probability distributions, Ref. [61] showed that such form of stringent, device-independent certification is infeasible for most prominent quantum supremacy distributions, requiring exponentially many samples.

In reality, however, the experimenters do have prior knowledge about the functioning of various components of their device and the model of the physical noise. This prompts one to move away from the minimal assumptions. If one insists on only making use of the classical output statistics, benchmarking protocols of the cross-entropy type [13,14,25], allow one to rule out distributions that are otherwise ideal but are corrupted by depolarizing noise with few samples. (Reference [14] derives an entropic condition on distributions that are ruled out by the cross entropy difference proposed in Ref. [13].) Thus, some amount of effort in the analysis of Google's experiment [25] is dedicated to benchmarking the error model and validating their assumptions. The downside is that this type of measure requires one to actually compute the ideal quantum supremacy probabilities, which are presumed to be extremely hard to compute when the number of qubits is large. For boson sampling, a weaker form of certification based on state discrimination [62] can be made fully efficient (that is, efficient in both the number of samples and computation time), but can only certify against a fixed adversarial distribution (for example, the uniform distribution).

Assuming the ability to change input states or measurement settings (trusted preparations and measurements), direct certifications for several quantum advantage architectures can be devised that are fully efficient. One such line of works [24,26,63] is based on the idea of fidelity witness of output states.

In our work, we offer a way to characterize FLO circuits assuming the ability to prepare certain product input states and perform trusted Pauli measurements. The certification is indirect as it certifies only a component of the experiment (the circuit) but not the quantum supremacy distribution itself. However, the protocol suggests that a direct certification of fermion sampling can be developed similar to that of boson sampling as our scheme is analogous to the characterization of linear optical networks [60] whereas no such scheme exists at present for random circuit sampling.

4. Relation to fermionic quantum computation

Quantum computing with (active) FLO circuits has received significant attention over the years. While FLO circuits with unentangled input states and measurements are efficiently simulable classically, they constitute a "maximally classical" subset of quantum circuits in the sense that an addition of any non-FLO unitary allows one to reach any unitary on the relevant Hilbert space [64]. Thus, similar to Clifford circuits, FLO circuits with additional resources constitute an interesting model of universal quantum computing [55,58]. Here we review the most important results about computational power of FLO circuits and their extensions. Common notions of simulation in the literature fall into two classes of *strong* and *weak* simulations: strong simulation refers to the ability to compute the marginal probability of any chosen outcome, whereas weak simulation refers to the ability to sample from the output probability distribution. Strong classical simulability of FLO circuits can be traced back to the work of Valiant [35] in which he introduced so-called matchgates for the purpose of studying algorithms for graphs. Assuming computational-basis input states and measurements, circuits of nearest-neighbor (NN) matchgates in one-dimensional (1D) layout can be strongly simulated on a classical computer in polynomial time, even with adaptive measurement in the computational basis [33]. Soon after the introduction of matchgates, their classical simulability was connected to exact solvability in physics as NN matchgate circuits can be mapped to evolutions of noninteracting fermions via the Jordan-Wigner transformation [33,34] and extended to classical simulability of dissipative FLO and nonunitary matchgates [65,66]. The geometric locality restriction is nontrivial as matchgate computation becomes quantum universal when the NN condition is lifted [67] or the linear chain of qubits is replaced by more general graphs [68].

When one considers arbitrary product input states, adaptive computation using FLO circuits with such inputs can be simulated classically [69]. This is in striking contrast to the case of Clifford circuits in which supplying single-qubit magic states and adaptive measurement in the computational basis suffices for universal quantum computation [70]. Since every fermionic state (or qubit state with a fixed parity) of fewer than four qubits is Gaussian [71,72], FLO circuits with computational-basis measurement must be supplied with at least a four-qubit magic input state to attain universality. The first example of such a state is $|a_8\rangle = 1/\sqrt{2}(|0000\rangle + |1111\rangle)$ (which can be converted to $|\Psi_4\rangle$ used in our scheme, see the proof of Lemma 15) was introduced in Ref. [58] along with the corresponding state-injection scheme for universal quantum computation using Ising anyons. A much more general result was established in Ref. [59] where it was showed that all non-Gaussian states, when supplied in multiple copies, allow one to perform universal quantum computation. Finally, weak classical simulability of FLO circuits with noisy magic input states was studied in Refs. [72,73],

Alternatively to magic input states, adding an arbitrary non-FLO gate [55,64,74] (see also [75]), or entangled measurements such as nondestructive parity measurement [55,76], also allows one to perform universal quantum computation. When the final measurement is restricted to only one qubit line and no adaptive measurement is allowed during the computation, the circuits are classically simulable in the strong sense even with magic input states. This result was first proven for an arbitrary product input state in Ref. [67] and observed to generalize to any product of $O(\log m)$-qubit states in Ref. [77]. Recently comprehensive investigation of the complexity landscape of FLO circuits with auxiliary resources are given in Ref. [77], which investigated the hardness of FLO circuits depending on the following: (i) whether the input is a product state or copies of entangled magic states, (ii) whether adaptive measurements are allowed, (iii) whether the final measurement is performed only on a single qubit or on all qubits. It was established there that strong simulation of FLO circuits is $\#P$-hard in all cases considered except ones that are already known to be classically simulable [33,67,69]. Using the standard postselection argument [11], it is possible to show that weak simulation FLO circuits with magic input states and no adaptive measurement implies collapse of the polynomial hierarchy (see Appendix G). This scenario coincides with one of the settings considered in this work. However, in this work we are concerned with establishing hardness of fermion sampling up to additive error, which, as explained earlier, is a property much harder to establish.

Organization of the paper.—First, in Sec. II we lay out basic notations and concepts, focusing mostly on the fermionic context. Then in Sec. III we formally define our quantum advantage proposal and give a high-level overview of our results and their significance. We also present there arguments in favor of experimental feasibility of our scheme. In Sec. IV we discuss possible applications of our work and present future research directions. In the subsequent Sec. V we prove that output probabilities of FLO circuits initialized in suitable magic states anticoncentrate for generic active and passive FLO circuits. These results, together with known [56] worst-case #P-hardness of probability distributions, is then used in Sec. VI to prove hardness of approximate fermion sampling. Section VII is devoted to the quantitative analysis of the Cayley path transformation for unitary and orthogonal groups. Section D focuses on polynomials associated to the probabilities in our FLO sampling scheme. In Sec. VIII we use technical results from the two preceding parts to prove, following Ref. [15] worst-to-average-case reduction for hardness of computing probabilities in our quantum advantage scheme. In the final Sec. IX we show that in the Jordan-Wigner encoding an unknown FLO unitary can be efficiently certified using resources scaling polynomially with the number of fermionic modes. The Appendix consists of four parts and contains auxiliary technical results. In Appendix A, we describe in detail the decomposition of FLO circuits into elementary one- and two-qubit gates. In Appendix E, we give details of the computations needed in the proof of anticoncentration of our results. In Appendix F, we prove a lemma concerning the stability of the FLO representations (an analog of the stability result proved standard boson sampling [78]), which is used in the tomography scheme of FLO unitaries. In Appendix C we prove the technical lemma concerning total variation distance between Haar measures on groups $G = \mathrm{U}(d), \mathrm{SO}(2d)$ and their deformations via Cayley path. Finally, in Appendix G we prove #P-hardness of probabilities in shallow-depth active FLO circuits.

II. NOTATION AND BASIC CONCEPTS

In this section we describe the main concepts and notation needed in the paper. Specifically, we introduce the language of second quantization, vital for describing fermionic systems. We define passive and active fermionic linear optical circuits. Finally, we survey Jordan-Wigner transformation, which allows implementation of fermionic systems and associated unitaries acting on them in terms of spin systems and standard quantum circuits. All these ingredients allow us to formally define our scheme for attaining quantum computational advantage with FLO circuits.

Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Normalized vectors in this space are denoted by $|\Psi\rangle, |\Phi\rangle$ etc. Such normalized vectors give rise to pure states, i.e., rank-1 non-negative operators on $\mathcal{H}$. For the sake of brevity, we use the notation $\Psi = |\Psi\rangle\langle\Psi|, \Phi = |\Phi\rangle\langle\Phi|$ etc. We use the symbol $\mathcal{D}(\mathcal{H})$ for the set of all (possibly mixed) quantum states on $\mathcal{H}$. Finally, by $\mathrm{U}(\mathcal{H})$ we denote group of unitary operators on $\mathcal{H}$. We consider a system of fermions with single-particle Hilbert space being $\mathbb{C}^d$. The Hilbert space associated to this system is a d-mode Fock space

$$\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d) = \bigoplus_{n=0}^{d} \bigwedge^{n}(\mathbb{C}^d), \tag{1}$$

where $\bigwedge^n(\mathbb{C}^d)$, is the totally antisymmetric subspace of $(\mathbb{C}^d)^{\otimes n}$ describing states consisting of exactly n fermions, and $\bigwedge^0(\mathbb{C}^d) = \mathrm{span}_{\mathbb{C}}(|0_F\rangle)$, where $|0_F\rangle$ is the Fock vacuum. Any basis $\{|1\rangle, |2\rangle, \ldots, |d\rangle\}$ of single-particle Hilbert space defines a family of creation and annihilation operators acting on $\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$: $f_j^\dagger$ and f_j, respectively, where $j = 1, 2, \ldots, d$. These operators satisfy canonical anticommutation relations $\{f_j, f_k^\dagger\} \equiv f_j f_k^\dagger + f_k^\dagger f_j = \delta_{j,k}$ and $\{f_j, f_k\} = \{f_j^\dagger, f_k^\dagger\} = 0$, with $\delta_{j,k}$ being the Kronecker symbol.

Given this set of creation and annihilation operators, it is natural to introduce the so-called Fock basis states, which forms a basis of $\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$, as

$$|\mathbf{x}\rangle := (f_1^\dagger)^{x_1} (f_2^\dagger)^{x_2} \cdots (f_d^\dagger)^{x_d} |0_F\rangle \tag{2}$$

for any $\mathbf{x} \in \{0, 1\}^d$, where we use the notation $(f_1^\dagger)^0 = \mathbb{I}$. Throughout the paper, we denote the set $\{1, \ldots, d\}$ as $[d]$. Given an arbitrary subset $\mathcal{X} \subset [d]$, it is also be useful to introduce the notation $|\mathcal{X}\rangle$ for the Fock basis state $|\mathbf{x}\rangle$ with $x_j = 1$ if $j \in \mathcal{X}$ and $x_j = 0$ otherwise. We also use $\binom{\mathcal{X}}{k}$ to denote the collection of subsets of finite set $\mathcal{X}$ of size k [we assume the convention $\binom{\mathcal{X}}{k} = \emptyset$ if $|\mathcal{X}| < k$].

Considering the direct sum decomposition of the Fock space into fixed particle-number subspaces in Eq. (1), a specific Fock basis state $|\mathcal{X}\rangle$ (with $|\mathcal{X}| = n$) is an element of the n-particle subspace $\bigwedge^n(\mathbb{C}^d)$. Note that since $\bigwedge^n(\mathbb{C}^d)$ can be regarded as a subspace of $(\mathbb{C}^d)^{\otimes n}$, it is natural to consider such a Fock basis state $|\mathcal{X}\rangle$ (where $\mathcal{X} = \{a_1, \ldots a_n\}$ with $a_i < a_j$ if $i < j$) as an element in $(\mathbb{C}^d)^{\otimes n}$, which is given by the formula

$$\begin{aligned} |\mathcal{X}\rangle &= |a_1\rangle \wedge |a_2\rangle \wedge \ldots \wedge |a_n\rangle \\ &= \frac{1}{\sqrt{n!}} \sum_{i_1, \ldots i_n = 1}^{n} \epsilon_{i_1, i_2, \ldots, i_n} |a_{i_1}\rangle \otimes |a_{i_2}\rangle \otimes \cdots |a_{i_n}\rangle. \end{aligned} \tag{3}$$

Here and throughout the paper we use the generalized Levi-Civita symbol, i.e., for any string of positive integers $(k_1, k_2, \ldots, k_n)$ with $k_i \neq k_j$ if $i \neq j$ we define $\epsilon_{k_1, k_2, \ldots, k_n} = (-1)^p$, where p is the parity of the permutation π for which $k_{\pi(i)} < k_{\pi(j)}$ if $i < j$ and p is its parity, while $\epsilon_{k_1, k_2, \ldots, k_n} = 0$ if some of the entries in $(k_1, k_2, \ldots, k_n)$ coincide.

A *passive fermionic linear optical transformation* on the n-particle subspace $\bigwedge^n(\mathbb{C}^d) \subset (\mathbb{C}^d)^{\otimes n}$ is given as a transformation $U^{\otimes n}$ restricted from being a $(\mathbb{C}^d)^{\otimes n} \to (\mathbb{C}^d)^{\otimes n}$ function to being a $\bigwedge^n(\mathbb{C}^d) \to \bigwedge^n(\mathbb{C}^d)$ map. Passive FLO can be understood abstractly as the irreducible representation of the low-dimensional symmetry group $\mathrm{U}(d)$ in the Hilbert space $\bigwedge^n(\mathbb{C}^d)$

$$\Pi_{\mathrm{pas}} : \mathrm{U}(d) \longrightarrow \mathrm{U}\left(\bigwedge^n(\mathbb{C}^d)\right), \tag{4}$$

$$U \longmapsto U^{\otimes n}\big|_{\bigwedge^n(\mathbb{C}^d)} . \tag{5}$$

That is, we get a representation of $\mathrm{U}(d)$ on a fixed particle fermionic subspace. A useful equivalent definition is that for any $U = e^{iK} \in \mathrm{SU}(d)$, Π_{pas} is the restriction of the Fock state unitary $e^{i/2\sum_{n,m} K_{nm} f_n^{\dagger} f_m}$ to the subspace $\bigwedge^n(\mathbb{C}^d)$.

An important concept when discussing passive FLO transformations are Slater determinant states. These are states of the form $|\Psi\rangle = |\xi_1\rangle| \wedge \xi_2\rangle \wedge \ldots \wedge |\xi_n\rangle$, where $\{|\xi_i\rangle\}_{i=1}^n \subset \mathbb{C}^d$ is a set of orthonormal vectors of the one-particle Hilbert space $\mathbb{C}^d$. By definition, Fock basis states are special cases of Slater determinant states. And passive FLO transformations act transitively on the set of Slater determinant states. The overlap between any two Slater determinant states, $|\Psi\rangle = |\xi_1\rangle| \wedge |\xi_2\rangle \wedge \ldots \wedge |\xi_n\rangle$ and $|\Phi\rangle = |\phi_1\rangle \wedge |\phi_2\rangle \wedge \ldots \wedge |\phi_n\rangle$, can be expressed by the simple determinant formula

$$\langle\Psi|\Phi\rangle = \det C,\ C_{i,j} = \langle\xi_i|\phi_j\rangle. \tag{6}$$

A standard way to measure fermionic systems is to perform particle-number measurement, i.e., a projective measurement in the Fock basis basis $|\mathbf{x}\rangle$ defined previously. Upon obtaining measurement result labeled by $\mathcal{X}$, numbers x_i have the interpretation of number of particles in mode i.

Let us next introduce the self-adjoint Majorana mode operators

$$m_{2j-1} = f_j + f_j^{\dagger}, \quad m_{2j} = -i\,(f_j - f_j^{\dagger}), \tag{7}$$

with anticommutation relations $\{m_j, m_k\} = 2\ \mathbb{I}\delta_{j,k}$. These operators define parity operator $Q = i^d \prod_{i=1}^{2d} m_i$ in $\mathcal{H}_{\mathrm{Fock}}$ (C^d). The subspace of $\mathcal{H}_{\mathrm{Fock}}(C^d)$ that corresponds to eigenvalue of $+1$ of Q is spanned by Fock states $|\mathbf{n}\rangle$ having even number of particles. In what follows, we refer to this vector space as the positive parity subspace and denote it by $\mathcal{H}^+_{\mathrm{Fock}}(\mathbb{C}^d)$. Majorana operators allow to define also active fermionic linear optical transformations. We say that a fermionic unitary U is *free*, *Gaussian*, or *linear optical*, if it can be written as an exponential of a quadratic Hamiltonian, i.e., $U = \mathrm{e}^{iH}$, where

$$H = \frac{i}{4}\sum_{j,k=1}^{2d} A_{j,k}\, m_j\, m_k, \tag{8}$$

and $A = -A^T \in \mathbb{R}^{2d\times 2d}$. Active FLO transformation form a group, which can be conveniently understood in terms of (projective) representation of the $\mathrm{SO}(2d)$ group [79]:

$$\Pi_{\mathrm{act}} : \mathrm{SO}(2d) \longrightarrow \mathrm{U}[\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)] \tag{9}$$

$$O \longmapsto \ \exp\left(\frac{1}{4}\sum_{i,j=1}^{2d} [\log(O)]_{ij}\ m_i m_j\right), \tag{10}$$

Often the restriction of Π_{act} to the positive parity subspace $\mathcal{H}^+_{\mathrm{Fock}}(\mathbb{C}^d)$ is considered, we do not use a new symbol for this restricted representation rather simply refer to it by writing $\Pi_{\mathrm{act}} : \mathrm{SO}(2d) \to \mathrm{U}[\mathcal{H}^+_{\mathrm{Fock}}(\mathbb{C}^d)]$. Pure positive-parity Gaussian states are defined as pure states the form $\Psi = \Pi(O)\,|0_F\rangle\langle 0_F|\,\Pi(O)^{\dagger}$, for $O \in \mathrm{SO}(2d)$. In other words, pure positive-parity Gaussian fermionic states are states that can be generated from the vacuum by active FLO transformations. Similarly, it is possible to define negative-parity pure fermionic Gaussian states as states generated by active FLO from, say, a Fock state with a single excitation.

If we look at the action on the operators, we get an actual (i.e., nonprojective) representation. In particular, a single Majorana operator evolves under an active FLO transformation as follows:

$$U^{\dagger}\ m_j\ U = \sum_{k=1}^{2d} O_{jk}\ m_k, \tag{11}$$

where $U = e^{-iHt}$ with $H = i/4\sum_{j,k=1}^{2d} A_{j,k} m_j m_k$ and $O = e^{-A} \in \mathrm{SO}(2d)$.

We also use the notation $\mathcal{G}_{\mathrm{pas}}$ and $\mathcal{G}_{\mathrm{act}}$ to denote, respectively, passive and active fermionic linear optical gates. The name comes from the fact that these gates transform single creation and Majorana operators to linear combination of creation and Majorana operators, respectively.

An important ingredient when discussing how to implement FLO transformations on qubit systems is the Jordan-Wigner transformation, that provides an equivalence between fermion and qubit systems through the unitary mapping $\mathcal{V}_{\mathrm{JW}} : \mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d) \to (\mathbb{C}^2)^{\otimes d}$ given as the following mapping between

$$\mathcal{V}_{\mathrm{JW}}\left((f_1^{\dagger})^{x_1}(f_2^{\dagger})^{x_2}\cdots(f_d^{\dagger})^{x_d}|0_F\rangle\right) = \bigotimes_{p=1}^{d} |x_p\rangle \tag{12}$$

for all $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \{0,1\}^d$, which in turn induces an isomorphic mapping between Majorana and spin operators

$$m_{2p-1} \mapsto \mathcal{V}_{\mathrm{JW}}\, m_{2p-1}\, \mathcal{V}_{\mathrm{JW}}^{\dagger} = Z_1 \cdots Z_{p-1} X_p, \tag{13}$$

$$m_{2p} \mapsto \mathcal{V}_{\mathrm{JW}}\, m_{2p}\, \mathcal{V}_{\mathrm{JW}}^{\dagger} = Z_1 \cdots Z_{p-1} Y_p, \tag{14}$$

where $p \in [d]$. To make the connection between fermions and qubit systems even more transparent, one often introduces the *occupation number notation* for vectors in

$\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$ as $|\mathbf{x}\rangle := (f_1^\dagger)^{x_1}(f_2^\dagger)^{x_2}\cdots f_d^\dagger)^{x_d}|0_F\rangle$ for any $\mathbf{x}\in\{0,1\}^d$. As the $|\mathbf{x}\rangle$ vectors are mapped via the Jordan-Wigner transformation to the computational basis states, they are also called the *fermionic computational basis states* in $\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$.

Since groups $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ are compact groups (for comprehensive introduction to the theory of Lie groups and their representations, see Refs. [80,81]), each possesses a unique normalized integration measure invariant under any group translation called Haar measure. We donate this measure by μ_G for the G one of the symmetry groups above. Invariance of μ_G means that any measurable subset $A\subset G$ and any $h\in G$, we have that

$$\mu(hA)=\mu(Ah)=\mu(A). \qquad (15)$$

The above condition to the level of expectation values (averages) reads

$$\int_G d\mu(g)f(gh)=\int_G d\mu(g)f(hg)=\int_G d\mu(g)f(g), \quad (16)$$

where f is any integrable function on G and $h\in G$. We denote by ν_{pas} the distribution of the unitaries $V=\Pi_{\mathrm{pas}}(U)$, where $\mathrm{U}\sim\mu_{\mathrm{U}(d)}$ and by ν_{act} distribution of the unitaries $\Pi_{\mathrm{act}}(O)$, where $O\sim\mu_{\mathrm{SO}(2d)}$. In order to keep the notation compact we suppress the dependence of these measures on d and n (values of these parameters are implied from the context).

Finally, we use the following notation to denote growth of functions: let f and g be two positive-valued functions. We write $f=O(g)$ if and only if $\lim_{x\to\infty}f(x)/g(x)<\infty$ and $f=o(g)$ if and only if $\lim_{x\to\infty}f(x)/g(x)=0$.

III. MAIN RESULTS

In this part we formally define our scheme for demonstration of quantum computational advantage and present informally the main results of this work. In the end we comment on the practical feasibility of our quantum advantage scheme.

Having reviewed the basic concepts needed, we are now ready to formally introduce our quantum advantage proposal, which is illustrated in Fig. 1. We have a system of $d=4N$ fermionic modes. The input state of the scheme is an N-fold tensor product the non-Gaussian magic state $|\Psi_4\rangle=(|1100\rangle+|0011\rangle)/\sqrt{2}$, i.e.,

$$|\Psi_{\mathrm{in}}\rangle=|\Psi_4\rangle^{\otimes N}. \qquad (17)$$

Note that states equivalent to this have been used in FLO computation schemes in Refs. [56,58,59]. After the initialization, a generic FLO operation is applied either respecting the particle-number conservation (passive scheme) or not (active scheme). Any FLO unitary can be decomposed into two-qubit FLO gates of linear depth either in diamond, triangle, or brickwall layouts, see Figs. 1 and 4. The choice of the FLO operation V is done via the probability distributions ν_{pas} and ν_{act} induced from the Haar measures on $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$, respectively (see Sec. II).

For a particular type of FLO circuit, the computational task we address is the ability to sample from the output distribution

$$p_{\mathbf{x}}(V,\Psi_{\mathrm{in}})=|\langle\mathbf{x}|V|\Psi_{\mathrm{in}}\rangle|^2, \qquad (18)$$

where output bitstring satisfies $|\mathbf{x}|=2N$ and $|\mathbf{x}|$-even for passive FLO and active FLO, respectively. This computational task is referred to as *fermion sampling*. We prove four main technical results that underpin the hardness of fermion sampling.

The first result is anticoncentration for FLO circuits. Informally speaking it states that for the considered family of circuits and fixed output $\mathbf{x}$ values $|\langle\mathbf{x}|V|\Psi_{\mathrm{in}}\rangle|^2$ are typically not much smaller compared to they average average value.

Result 1: (Anticoncentration for generic FLO circuits.) *Let $\nu=\nu_{\mathrm{pas}}$ or $\nu=\nu_{\mathrm{act}}$ be uniform distribution over passive and, respectively, active FLO circuit acting on $4N$ fermion modes. Let Ψ_{in} be the input state to our quantum advantage proposal. Then, there exist a constant $C\geq 1$ such that for every outcome* $\mathbf{x}$ *and for every $\alpha\in[0,1]$*

$$\Pr_{V\sim\nu}\left(p_{\mathbf{x}}(V,\Psi_{\mathrm{in}})>\frac{\alpha}{|\mathcal{H}|}\right)>\frac{(1-\alpha)^2}{C}, \qquad (19)$$

where $\mathcal{H}=\bigwedge^{2N}(\mathbb{C}^{4N})$ for passive FLO and $\mathcal{H}=\mathcal{H}^+_{\mathrm{Fock}}(\mathbb{C}^{4N})$ for active FLO.

The formal version of this result is given in Theorem 1. It is important to emphasize that in the course of the proof of this results we do not use the property of gate sets of interest forming an (approximate) 2-design [26]. In fact, it can be proved that measures $\nu_{\mathrm{pas}},\nu_{\mathrm{act}}$ do not form a projective 2-design. We perform the proof of anticoncentration by heavily using group-theoretical techniques and particular properties of fermionic representations of symmetry groups $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$.

In line with standard methodology based on Stockmeyer's algorithm [53], anticoncentration, and hiding property we reduce approximate sampling from $\{p_{\mathbf{x}}(V,\Psi_{\mathrm{in}})\}$ to approximate computation of particular probability $p_{\mathbf{x}_0}(V,\Psi_{\mathrm{in}})$ (see Theorem 2). This allows us to prove hardness of approximate fermion sampling in Theorem 3 by conjecturing noncollapse of the polynomial hierarchy (cf. Conjecture 2) and average-case hardness of computation approximate computation of $p_{\mathbf{x}_0}(V,\Psi_{\mathrm{in}})$ in relative error (cf. Conjecture 1).

Remark 1: It is important to stress that in the passive FLO case our anticoncentration results *do not* follow from anticoncentration results for the determinant proved in Ref. [10]. The reason is that our probability amplitudes can only be expressed via determinants of submatrices of $U \in \mathrm{U}(d)$ (cf. Sec. D). Also, these submatrices cannot be approximated via Gaussian matrices (we work in the regime in which the number of modes d is comparable to the total number of particles n).

To give evidence for Conjecture 1, we prove a worst-to-average-case reduction that allows us to prove a weaker version of approximate hardness result.

Result 2: (Worst-case to average reduction for approximate computation of probabilities.) *Let $\nu = \nu_{\mathrm{pas}}$ or $\nu = \nu_{\mathrm{act}}$ be uniform distribution over passive and, respectively, active FLO circuit acting on $4N$ fermion modes, and let Ψ_{in} be the input state to our quantum advantage proposal. It is #P-hard to approximate probability $p_{\mathbf{x}_0}(V, \Psi_{\mathrm{in}})$ to within accuracy $\epsilon = \exp[-\Theta(N^6)]$ with probability greater than $1 - o(N^{-2})$ over the choice of $V \sim \nu$.*

Formal versions of above result can be found in Theorem 7. To obtain the above result we generalize the method developed recently by Movassagh [15] in the context of random quantum circuits. The key technical ingredient a Cayley path, which gives rational interpolation between quantum circuits. We realize that, for the purpose of the two reductions given above, it is possible to apply it directly on one level of the Lie group underlying a particular class of FLO transformation [$\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ for passive and active FLO, respectively]. We then use the fact that fermionic representations can be realized low degree of polynomials in entries of matrices of appropriate symmetry groups. This observation allows us to adapt the reduction method of Movassagh with relative ease.

Remark 2: In the course of the proof of the above result we realize a technical issue in Movassagh [15]. Correction of the proof gives worse than claimed tolerance for error $\epsilon = \exp[-\theta(N^{4.5})]$ (for the Google layout), which is still better than the one claimed here. On the other hand, application of our reduction in conjunction with recent improvements over the Paturi lemma by Bouland *et al.* and Kondo *et al.* [31,82] (published after the completion of this work) boost error tolerance of our scheme to $\epsilon = \exp\{-\theta[N^2 \log(N)]\}$.

Finally, the experimental feasibility of our proposal is further increased by the fact that due to the structure of FLO circuits, they can be efficiently certified using resources scaling polynomially with the system size.

Result 3: (Efficient tomography of FLO circuits.) *Let V be an unknown active FLO circuit on a system of d qubits that encodes d fermionic modes. Assume we have access to computational basis measurements and single-qubit gates. Then V can be estimated up to accuracy ϵ in the diamond norm by repeating $r \approx d^3/\epsilon^2$ rounds of experiments, each involving $O(d^2)$ independent single-qubit state preparations and single Pauli measurements at the end of the circuit.*

The rigorous formulation of the above, together with the explicit protocol for carrying out tomography, is provided in Sec. IX. Importantly, our method avoids exponential scaling inherent to the general multiqubit tomography protocols. Moreover, it can also be viewed as a fermionic analog of the certification methods developed previously in the context of photonics and bosonic linear optics [83–85].

Implementation of the scheme

It is important to stress that our proposal has a strong potential for experimental realization, e.g., on quantum processors with superconducting qubit architectures. The actual implementation should be feasible already on near-term quantum devices, as the construction of parametric

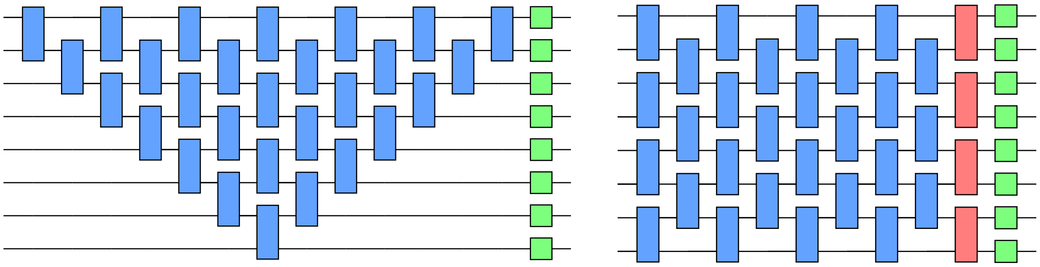


FIG. 4. Circuit layouts implementing arbitrary passive and active FLO transformations. These layouts are based on the decomposition of arbitrary elements of the $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ groups into a sequence of nearest-neighbor Givens rotations and a diagonal matrix. The depicted two-qubit gates in the passive FLO case are of the type $D_{\mathrm{pas}}(\alpha_1, \alpha_2)$ [see Eq. (20)], while the single-qubit gates are Z rotations. The two-qubit gates in the active FLO case are of the type $D_{\mathrm{act}}(\{\beta_i\})$ [see Eq. (21)] and the single-qubit gates are Pauli unitaries. The extra layer of red colored two-qubit gates are only needed in the active case. The decomposition of the two-qubit gates $D_{\mathrm{pas}}(\alpha_1, \alpha_2)$ and $D_{\mathrm{act}}(\{\beta_i\})$ into native gates of superconducting qubit architectures are provided in Fig. 5.

programmable passive linear optical circuits, due to their relevance in quantum chemistry, has been already experimentally demonstrated on Google's Sycamore quantum processor [37].

The preparation of the input fermionic magic state $|\Psi_4\rangle^{\otimes N}$, vital to our proposal, can be performed by applying on the computational basis state $|0\rangle^{\otimes 4N}$ a simple constant depth circuit consisting of three CNOTs and three one-qubit gates per quadruple blocks of qubits as shown in Fig. 1. One can implement an arbitrary passive FLO (or *basis rotation* in the quantum chemistry lingo) in linear depth using only nearest-neighbor gates and assuming a minimal linearly connected architecture [38,39]. Two such layouts are depicted in Fig. 4. In terms of two-qubit gates, the triangle layout has a depth of $d-1$, while the depth of the brickwall layout is only $d/2$. These circuits are analogous to the layouts of boson sampling circuits [5,6] and are based on decomposing a unitary $U \in \mathrm{U}(d)$ into individual Givens rotations, which we describe in Appendix A. In the passive FLO case, the two-qubit gates have the form

$$D_{\mathrm{pas}}(\alpha_1,\alpha_2) = (e^{-i\alpha_1 Z_1/2} e^{i\alpha_1 Z_2/2})\, e^{i\alpha_2(X_1X_2+Y_1Y_2)/2}, \quad (20)$$

and the final one-qubit gates are Z rotations. The triangle and the brickwall layout can also be used to decompose an arbitrary active FLO operation [39,40], however, in this case the two-qubit gates will have a more complicated structure (as they arise from merging several Givens rotations):

$$D_{\mathrm{act}}(\{\beta_i\}) = (e^{i\beta_5 Z_1/2} e^{i\beta_6 Z_2/2})\, e^{i(\beta_3 X_1X_2+\beta_4 Y_1Y_2)/2} \times (e^{i\beta_1 Z_1/2} e^{i\beta_2 Z_2/2}), \quad (21)$$

and the one-qubit unitaries at the end of the circuit are either Pauli matrices or identities. The derivations of these statements are given in Appendix A, they are based on the decomposition of arbitrary elements of the $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ groups into a sequence of nearest-neighbor Givens rotations and a diagonal matrix. The passive FLO representation of the Givens rotations and of the diagonal matrix are then translated to two-qubit gates of the type $D_{\mathrm{pas}}(\alpha_1,\alpha_2)$ and a series single-qubit Z rotations in a layout the depicted Fig. 4. In the active FLO case several represented Givens rotations are merged into two-qubit gates of type $D_{\mathrm{act}}(\{\beta_i\})$ and the single-qubit unitaries at the end of the circuit are Pauli gates.

In the experimental demonstration of programmable passive FLO transformations by the Google team [37], the native gates of the Sycamore processors, the $\sqrt{i\mathrm{SWAP}}$ gates and single-qubit Z rotations, are used. The iSWAP and $\sqrt{i\mathrm{SWAP}}$ gates, defined as

$$i\mathrm{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\sqrt{i\mathrm{SWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (22)$$

are exactly introduced in quantum computing as standard gates because they are native in superconducting qubit architectures [86]. It is important to note that these gates are actually FLO gates. This lucky coincidence supports the feasibility of our proposal, since the Givens rotations for passive FLO, i.e., the two-qubit gates $D_{\mathrm{pas}}(\alpha_1,\alpha_2)$ in the layouts of Fig. 4, can be decomposed into two $\sqrt{i\mathrm{SWAP}}$ gates and four single qubit Z-rotations as depicted in part (a) of Fig. 5. The two-qubit gates used in the active FLO setup, $D_{\mathrm{act}}(\{\beta_i\})$, can be decomposed into three iSWAP gates shown in part (b) of Fig. 5.

In Result 1 the anticoncentration of passive and active FLO circuits is presented. This result gives for active FLO random circuits an estimation of the constant C_{act} which is computed by bounding the expectation values from the Paley-Zygmund inequality $(\mathbb{E}X)^2/\mathbb{E}X^2$, with $X = |\langle \mathbf{x}| V |\Psi_{\mathrm{in}}\rangle|^2$ where V is the random circuit. As shown in Fig. 4, we have assumed that the random circuits have linear depth with respect to the number of modes, nonetheless it is still possible that the anticoncentration property is obtained for lower depths. Note that at shallow depth the so-called property of data hiding is also present, this property states that given $|\langle \mathbf{x}| V |\Psi_{\mathrm{in}}\rangle|^2$ and a fixed state $|\mathbf{x}_0\rangle$ there is a FLO circuit $V_{\mathbf{x}}$ such that $|\langle \mathbf{x}| V |\Psi_{\mathrm{in}}\rangle|^2 = |\langle \mathbf{x}_0| V_{\mathbf{x}} |\Psi_{\mathrm{in}}\rangle|^2$. For more detail of the data hiding property see Lemma 1. To test that anticoncentration is obtained at shallow depth, we simulate noiseless fermion sampling experiments with up to 28 qubits using the IBM Qiskit simulator [87]. As in Fig. 1, we initialize the input with a fixed number of quadruples $|\Psi_4\rangle$. We perform simulations for $1, 2, \ldots, 7$ quadruples, for each of these, we compute the output probability of a fixed output string for an average of 26 000 circuits for each number of quadruples. The circuits used have the brickwall architecture from Fig. 4. From each random circuit the output probability of a fixed output string is computed. With the obtained output probabilities, the expectation values in the Paley-Zygmund inequality $(\mathbb{E}X)^2/\mathbb{E}X^2$ are computed. In Fig. 2 we plot for each fixed number of quadruples the minimum depth required for the ratio of expectation values to surpass a certain threshold and compare to the case where the depth is linear with respect to the number of modes. The simulations suggest that the depths required to obtain

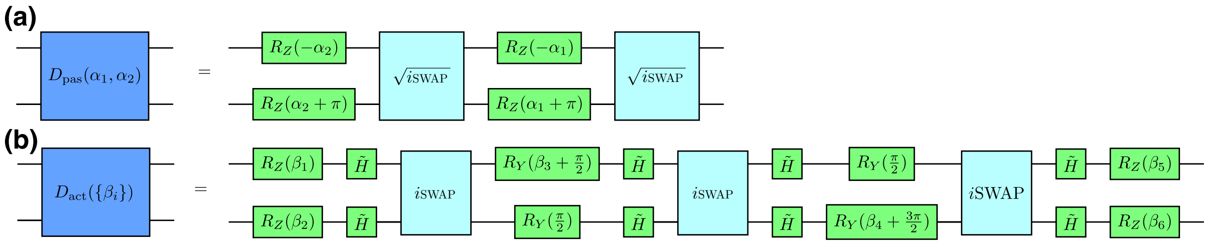


FIG. 5. Decomposition of (a) the Givens rotation in the passive FLO setting, $D_{\mathrm{pas}}(\alpha_1,\alpha_2)$, in terms of $\sqrt{i\mathrm{SWAP}}$ gates and (b) the merged Givens rotations in the active FLO setting, $D_{\mathrm{act}}(\{\beta_i\})$, in terms of $i\mathrm{SWAP}$ gates. $R_W(\alpha)$ (with $W\in\{X,Y,Z\}$) denotes the one-qubit rotation gate $e^{i\alpha W}$. The gates $\tilde{H}$ are defined by the relations $\tilde{H}Z\tilde{H}=Y$ and $\tilde{H}Y\tilde{H}=Z$.

anticoncentration are shallow as the number of modes is increased, wherein the case for seven initial quadruples and a threshold of 0.4 the number of layers required is 7. Moreover, in the random circuit sampling setup it has been found that anticoncentration is obtained in logarithmic depth, our simulations show that this could also be the case in our setup.

IV. DISCUSSION AND OPEN PROBLEMS

We believe that our results and techniques used to establish them will be of relevance also for problems not directly related to fermion sampling. The first group of potential applications is related to the structure of our quantum advantage scheme. As discussed in the introduction, quantum advantage proposals are typically not constructed because of their practical usefulness. However, recently several proposals for applications of boson sampling [10] and Gaussian boson sampling [22,23] have been suggested. These include combinatorial optimization problems [88,89], calculation of Franck-Condon profiles for vibronic spectra [90,91], molecular docking [92], and machine learning using graph kernels [93]. Admittedly a feasible scalability is unlikely for those applications that should produce as an answer one specific bitstring (or a few specific bitstrings), as the output probabilities are generically exponentially suppressed. However, different types of applications than those, like dense graph sampling or using graph samples for graph-kernel methods, might turn out to be useful. It should be mentioned that all of those application are based on the fact that this type of sampling, unlike the generic random circuit sampling, have a very structured nature, as they are based on the fact that one can sample with probabilities proportional to the permanents and Hafnians of certain matrices constructed from the circuit description. In particular, most of the mentioned applications use the fact that one can sample with probabilities proportional to the permanents and Hafnians of certain matrices constructed from the circuit description. These polynomial functions of matrix entries encode interesting properties, e.g., for adjacency matrices of graphs they provide the number of cycle covers and perfect matchings of the graph, respectively. In our proposal similar polynomials appear when describing the sampling probabilities: the mixed discriminant [57] and their generalizations (e.g., mixed Pfaffian [94]), which also encode important graph properties. Thus, we have good reasons to believe that our proposal, besides providing a robust computational advantage setup, might also be used for other algorithms with interesting applications.

We also want to emphasize the universality of our techniques for establishing anticoncentration and worst-to-average-case reduction for structured random circuits. Our anticoncentration result exploits group-theoretic properties of fermionic circuits and can be likely generalized to other scenarios where low-dimensional group structure appears. Moreover, our generalization of the worst-to-average-case reduction of Movassagh [15] can be applied to any sampling problem provided outcome probabilities can be interpreted as polynomials on low-dimensional groups (Cayley transformation that underlines the reduction can be defined on arbitrary Lie groups). For example, one can view boson sampling in the first quantization picture where the group of linear optical networks acts on the totally symmetric subspace of N qudits, where d is the number of modes (or other representation when bosons are partially distinguishable [95]), as opposed to the totally antisymmetric representation for fermions.

We conclude this part with stating a number of interesting problems that require further study.

Role of non-Gaussianity for hardness and anticoncentration. In practice magic states are not perfect and a high level of noise can bring these states into the convex hull of Gaussian states [58,72,73], which we know are efficiently simulable classically under FLO evolutions and Gaussian measurements. How much noise does our hardness result tolerate? The same question applies to anticoncentration of output probabilities, in which case we do not yet have a proof that FLO circuits with Gaussian inputs do not anticoncentrate, but we have numerical evidence that proofs based on the Paley-Zygmund inequality do not work (see Remark 6).

Fermion sampling with less magic. In our scheme, all the input qubit lines are injected with magic states. Do the hardness result and anticoncentration hold if we use only, say, $O(\log m)$ magic states?

Algorithms for classical simulation. Devising algorithms to approximately simulate FLO circuits with magic input states on average would not only lead to useful applications (for example, in the context of quantum chemistry), but is also vital to understand the complexity landscape of random FLO circuits. For the RCS scheme employed in Google's experiment with qubits placed on a 2D grid, advances in classical simulation techniques imposed a limit on the robustness of the average-case hardness that can be achieved with the current worst-to-average-case reduction that is agnostic to the circuit architecture and depth [96].

Tomography and certification of FLO circuits and fermion sampling. In this work, we gave only an efficient method to estimate an unknown FLO circuit (A related benchmarking of FLO circuits was recently proposed in Ref. [97]). It is interesting to extend our scheme beyond unitary circuits and to devise a method for which sample complexity and number of experimental settings exhibit an optimal scaling with the system size. Additionally, with further assumption on how the quantum device operates (e.g., the noise model), is there a simple diagnostic tool for fermion sampling similar to cross-entropy benchmarking for RCS [13]?

V. ANTICONCENTRATION OF FLO CIRCUITS

In this section we prove that outcome probabilities in fermionic circuits initialized in the state Ψ_{in} anticoncentrate for Haar random fermionic linear optical circuits. We prove anticoncentration for both passive and active fermionic linear optics. Our proof is based on interpretation of these circuits in terms of representation of group $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$, where d is the number of fermionic modes used.

Let $\mathcal{H}$ be a Hilbert space and let $\{|\mathbf{x}\rangle\}$ be a fixed (computational) basis of $\mathcal{H}$. For $V\in\mathrm{U}(\mathcal{H})$ and a pure state $|\Psi\rangle$. In what follows we denote by $p_{\mathbf{x}}(V,\Psi)$ the probability of obtaining outcome $\mathbf{x}$ on some input state $|\Psi\rangle$ on which unitary V was applied. Born rule implies

$$p_{\mathbf{x}}(V,\Psi)=|\langle\mathbf{x}|\,V|\Psi\rangle\,|^2. \tag{23}$$

In what follows we restrict our attention to $\mathcal{H}=\bigwedge^{2N}(\mathbb{C}^{4N})$ (for passive FLO) and $\mathcal{H}=\mathcal{H}^{+}_{\mathrm{Fock}}(\mathbb{C}^{4N})$ (for active FLO). Moreover, for $\mathbf{x}\in\{0,1\}^{4N}$ vectors $|\mathbf{x}\rangle$ denote standard Fock states (cf. Sec. II). In both of the cases considered the set of allowed $\mathbf{x}$ is different (see Theorem 1 for more details).

Definition 1: (Anticoncentration of ensemble of unitary matrices). Let ν be an ensemble (probability distribution) of unitary matrices $\mathrm{U}(\mathcal{H})$. We say that that ν exhibits anticoncentration on input state $|\Psi\rangle$ if and only if for every outcome $\mathbf{x}$ of computational basis measurement

$$\Pr_{V\sim\nu}\left(p_{\mathbf{x}}(V,\Psi)>\frac{\alpha}{|\mathcal{H}|}\right)>\beta, \tag{24}$$

where α,β are positive constants.

Remark 3: In this work we are concerned with families of probability distributions that are defined on Hilbert spaces of increasing dimension, parametrized by the total number of fermionic modes d. In this context, motivated by the structure of the proof of hardness of sampling (see Theorem 3) we are interested in cases when $\alpha,\beta=\Theta(1)$, i.e., are independent on $|\mathcal{H}|$.

Below we state our main result regarding anticoncentration of fermionic linear circuits initialized in the tensor product of Fermionic magic states

$$|\Psi_{\mathrm{in}}\rangle=|\Psi_4\rangle^{\otimes N}, \tag{25}$$

where $|\Psi_4\rangle=1/\sqrt{2}(|0011\rangle+|1100\rangle)$. Note that $|\Psi_{\mathrm{in}}\rangle\in\bigwedge^{2N}(\mathbb{C}^{4N})$.

Theorem 1: (Anticoncentration for fermionic linear optical circuits initialized in product of magic states.) *Let $\mathcal{H}_{\mathrm{pas}}=\bigwedge^{2N}(\mathbb{C}^{4N})$ and let $\mathcal{H}_{\mathrm{act}}=\mathcal{H}^{+}_{\mathrm{Fock}}(\mathbb{C}^{4N})$ be Hilbert spaces describing $2N$ Fermions in $4N$ modes and positive parity Fermions in $4N$ modes. Let $\mathcal{G}_{\mathrm{pas}}$ and $\mathcal{G}_{\mathrm{act}}$ be, respectively, passive and active FLO transformations acting on the respective Hilbert spaces and distributed according to the uniform measures ν_{pas} and ν_{act} (see Sec. II). Let $|\Psi_{\mathrm{in}}\rangle$ be the initial state to which both families of circuits are applied. Then, for every $\mathbf{x}$ of Hamming weight $|\mathbf{x}|=2N$ we have*

$$\Pr_{V\sim\nu_{\mathrm{pas}}}\left(p_{\mathbf{x}}(V,\Psi_{\mathrm{in}})>\frac{\alpha}{|\mathcal{H}_{\mathrm{pas}}|}\right)>\frac{(1-\alpha)^2}{C_{\mathrm{pas}}}, \tag{26}$$

where $C_{\mathrm{pas}}=5.7$ and $|\mathcal{H}_{\mathrm{pas}}|=\binom{4N}{2N}$. Moreover, for every $\mathbf{x}$ with even Hamming weight we have

$$\Pr_{V\sim\nu_{\mathrm{act}}}\left(p_{\mathbf{x}}(V,\Psi_{\mathrm{in}})>\frac{\alpha}{|\mathcal{H}_{\mathrm{act}}|}\right)>\frac{(1-\alpha)^2}{C_{\mathrm{act}}}, \tag{27}$$

where $C_{\mathrm{act}}=16.2$ and $|\mathcal{H}_{\mathrm{act}}|=2^{4N}/2$.

Proof. In order to prove Eqs. (26) and (26) we start with a standard tool used when proving anticoncentration—the Paley-Zygmund inequality. It states that for arbitrary nonnegative bounded random variable X and for $0<\alpha<1$, we have

$$\Pr_X(X > \alpha\mathbb{E}X) \geq (1-\alpha)^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}X^2}. \tag{28}$$

We use this bound for $X = |\langle \mathbf{x}| V |\Psi_{\text{in}}\rangle|^2$, where $V \sim \nu_{\text{pas}}$ or $V \sim \nu_{\text{pas}}$. Recall that, as explained in Sec. II, linear circuits $\mathcal{G}_{\text{pas}}$ and $\mathcal{G}_{\text{act}}$ can be understood in terms of representations of symmetry groups U(d) and SO($2d$). Haar measures on these symmetry groups induce uniform distributions on the $\mathcal{G}_{\text{pas}}$ and $\mathcal{G}_{\text{act}}$. Therefore, for $k = 1, 2$ we have

$$\mathop{\mathbb{E}}_{V\sim\nu}\left[p_{\mathbf{x}}(V,\Psi_{\text{in}})^k\right] = \int_G d\mu(g)\left[\text{tr}(|\mathbf{x}\rangle\langle\mathbf{x}|\,\Pi(g)\Psi_{\text{in}}\Pi(g)^\dagger)\right]^k, \tag{29}$$

where μ is the Haar measure on a Lie group G, and Π is a unitary representation of G in a suitable Hilbert space $\mathcal{H}$. The case of passive FLO corresponds to $G = \text{U}(4N)$ and $\Pi = \Pi_{\text{pas}}$ while for active FLO we have $G = \text{SO}(8N)$ and $\Pi = \Pi_{\text{act}}$ [cf. Eqs. (4) and (9)]. Both groups are irreducibly represented in Hilbert spaces $\mathcal{H}_{\text{act}}$ and $\mathcal{H}_{\text{pas}}$ by virtue of Schur lemma unitaries $\Pi(g)$ forming a 1-design. Consequently,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{V\sim\nu}\left[p_{\mathbf{x}}(V,\Psi_{\text{in}})\right] \\
&\quad= \int_G d\mu(g)\left[\text{tr}(|\mathbf{x}\rangle\langle\mathbf{x}|\,\Pi(g)\Psi_{\text{in}}\Pi(g)^\dagger)\right] \\
&\quad= \text{tr}\left(|\mathbf{x}\rangle\langle\mathbf{x}|\int_G d\mu(g)\left[\Pi(g)\Psi_{\text{in}}\Pi(g)^\dagger\right]\right) = \frac{1}{|\mathcal{H}|},
\end{aligned} \tag{30}$$

where in the last equality we use the 1-design property and the fact that $|\mathbf{x}\rangle \in \mathcal{H}$ is a normalized vector. Computation of the second moment can be greatly simplified by the usage of group theory. Let us first rewrite $\mathop{\mathbb{E}}_{V\sim\nu}\left[p_{\mathbf{x}}(V,\Psi_{\text{in}})^2\right]$ in the form convenient for computation:

$$\begin{aligned}
&\mathop{\mathbb{E}}_{V\sim\nu}\left[p_{\mathbf{x}}(V,\Psi_{\text{in}})^2\right] \\
&\quad= \int_G d\mu(g)\left[\text{tr}(|\mathbf{x}\rangle\langle\mathbf{x}|^{\otimes 2}\,\Pi(g)^{\otimes 2}\Psi_{\text{in}}^{\otimes 2}(\Pi(g)^\dagger)^{\otimes 2}\right] \\
&\quad= \text{tr}(A_{\Pi,G}\Psi_{\text{in}}\otimes\Psi_{\text{in}}),
\end{aligned} \tag{31}$$

where, due to unitarity of representation Π, and invariance of Haar measure under transformation $g \mapsto g^{-1}$

$$A_{\Pi,G} = \int_G d\mu(g)\left[\Pi(g)^{\otimes 2}\,|\mathbf{x}\rangle\langle\mathbf{x}|^{\otimes 2}\,(\Pi(g)^\dagger)^{\otimes 2}\right]. \tag{32}$$

Operator $A_{\Pi,G}$ acts on two copies of the original Hilbert space, $\mathcal{H}\otimes\mathcal{H}$ and is a manifestly G invariant in the sense that for all g we have $[A_{G,\Pi}, \Pi(g)^{\otimes 2}] = 0$. The integration in Eq. (32) can be carried out explicitly because objects in question have very specific properties that are rooted in the fact that Fock states constitute generalized coherent states of the considered representations of U($4N$) and SO($8N$) (cf. Remark 5 for more details). Let $|\Psi\rangle$ be a fixed pure state in $\mathcal{H}$ and let $|\mathbf{x}\rangle$ be a fixed fermionic Fock state belonging to appropriate Hilbert space $\mathcal{H}$

$$\begin{aligned}
&\exists g \in G \text{ such that} \\
&\quad \Psi = \Pi(g)\,|\mathbf{x}\rangle\langle\mathbf{x}|\,\Pi(g)^\dagger \iff |\Psi\rangle^{\otimes 2} \in \tilde{\mathcal{H}},
\end{aligned} \tag{33}$$

where $\tilde{\mathcal{H}} \subset \mathcal{H}\otimes\mathcal{H}$ is the carrier space of certain unique irreducible representation of G. In other words, it appears as one of the irreducible representations in the decomposition of the space $\mathcal{H}\otimes\mathcal{H}$, where G is represented via $g \mapsto \Pi(g)^{\otimes 2}$. Let $\tilde{\mathbb{P}}$ be the orthonormal projector onto $\tilde{\mathcal{H}} \subset \mathcal{H}\otimes\mathcal{H}$. Due to property (33) we get that $\text{supp}(A_{\Pi,G}) \subset \tilde{\mathcal{H}}$. Combining this with G-invariance of $A_{\Pi,G}$ we get, using Schur lemma, that $A_{\Pi,G}$ must be proportional to $\tilde{\mathbb{P}}$. The proportionality constant follows easily from normalization of $A_{\Pi,G}$. Putting this all together we obtain

$$A_{\Pi,G} = \frac{1}{|\tilde{\mathcal{H}}|}\tilde{\mathbb{P}}. \tag{34}$$

Inserting the expressions for the first and second moments to Payley-Zygmund inequality we get

$$\Pr_{V\sim\nu}\left(p_{\mathbf{x}}(V,\Psi_{\text{in}}) > \frac{\alpha}{|\mathcal{H}|}\right) \geq (1-\alpha)^2\frac{|\tilde{\mathcal{H}}|}{|\mathcal{H}|^2}\frac{1}{\text{tr}(\tilde{\mathbb{P}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})}. \tag{35}$$

From the above expression it is clear that anticoncentration is controlled by (i) the ratio of $|\tilde{\mathcal{H}}|/|\mathcal{H}|^2$ and (ii) expectation value $\text{tr}(\tilde{\mathbb{P}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})$. We give explicit forms of the projectors $\tilde{\mathbb{P}}$, as well as the dimensions $|\tilde{\mathcal{H}}|$ for both passive and active FLO in Lemmas 12 and 14 in Appendix E 3. From the expressions given there we obtain [98]

$$\frac{|\tilde{\mathcal{H}}_{\text{pas}}|}{|\mathcal{H}_{\text{pas}}|^2} \geq \frac{1}{N}, \quad \frac{|\tilde{\mathcal{H}}_{\text{act}}|}{|\mathcal{H}_{\text{act}}|^2} \geq \frac{1}{\sqrt{\pi N}}. \tag{36}$$

For passive FLO this gives [recall that $|\mathbf{x}| = 2N$ and $\mathcal{H}_{\text{pas}} = \bigwedge^{2N}(\mathbb{C}^{4N})$]

$$\begin{aligned}
&\Pr_{V\sim\nu_{\text{pas}}}\left(p_{\mathbf{x}}(V,\Psi_{\text{in}}) > \frac{\alpha}{\binom{4N}{2N}}\right) \\
&\quad\geq (1-\alpha)^2\frac{1}{N}\frac{1}{\text{tr}(\mathbb{P}_{\text{pas}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})}.
\end{aligned} \tag{37}$$

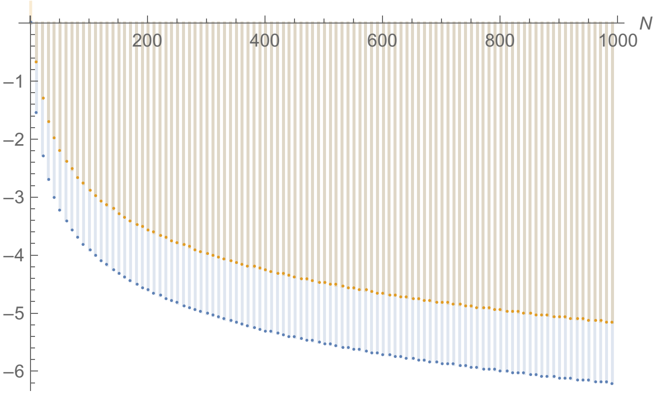


FIG. 6. Plots of the logarithm of the expression (E73) (blue) and $\log(C_{\text{pas}}/N) = \log(5.7/N)$ (orange), which constitutes a valid upper bound for all $N \leq 1000$.

Similarly, for active FLO we obtain [recall that $|\mathbf{x}|$ is even and $\mathcal{H}_{\text{act}} = \mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^{4N})$] we have

$$\Pr_{V\sim\nu_{\text{act}}}\left(p_{\mathbf{x}}(V,\Psi_{\text{in}}) > \frac{\alpha}{2^{4N-1}}\right) \geq (1-\alpha)^2 \frac{1}{\sqrt{\pi N}} \frac{1}{\operatorname{tr}(\mathbb{P}_{\text{act}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})}. \tag{38}$$

In order to complete the proof we need the following inequalities:

$$\operatorname{tr}(\mathbb{P}_{\text{pas}}\Psi_{\text{in}}\otimes\Psi_{\text{in}}) \leq \frac{C_{\text{pas}}}{N},\ \operatorname{tr}(\mathbb{P}_{\text{act}}\Psi_{\text{in}}\otimes\Psi_{\text{in}}) \leq \frac{C_{\text{act}}}{\sqrt{\pi N}}. \tag{39}$$

Proof of the above relies on the explicit form of the projectors as well as some combinatorial considerations. The details are given in the Appendix (see specifically Lemma 13 for of passive FLO and Lemma 15 for active FLO). ■

Remark 4: In the course of proving Eq. (39) in Appendices we arrive at upper bounds on $\operatorname{tr}(\mathbb{P}_{\text{pas}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})$ and $\operatorname{tr}(\mathbb{P}_{\text{act}}\Psi_{\text{in}}\otimes\Psi_{\text{in}})$ that are efficiently computable as a function of N. The numerics shown in Figs. 6 and 7 strongly suggest that the bounds provided by the values $C_{\text{pas}} = 5.7$ and $C_{\text{act}} = 16.2$ given here are not tight and can be improved by better proof techniques, specifically up to $C_{\text{pas}} \leq 2.4$ and $C_{\text{act}} \leq 2.7$.

Remark 5: The existence of projector $\tilde{\mathbb{P}}$ such that equivalence in Eq. (33) holds follows from the group-theoretical characterizations of Slater determinants as well as pure fermionic Gaussian states with positive parity. Namely, these classes of states constitute examples the so-called *generalized coherent states* of simple, compact, and connected Lie groups [SU(d) and spin($2d$)] that are irreducibly represented in the appropriate Hilbert space [$\bigwedge^n(\mathbb{C}^d)$ and $\mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^d)$, respectively]. The fact that such classes of states can be characterized via the quadratic condition $A\,|\Psi\rangle^{\otimes 2} = 0$ is a known result in algebraic geometry [99]. This was translated to the quantum information language in Ref. [100], later rephrased as Eq. (33) and used to characterize correlations in systems consisting of indistinguishable particles [101,102] (see also Chapter 3 of Ref. [103]). An equivalent characterization of pure fermionic Gaussian states was also independently discovered by Bravyi in the context of fermionic quantum information [65] (see also Ref. [72]).

Remark 6: A curious reader may wonder whether anticoncentration holds also if FLO circuits that are initialized in free Gaussian states Ψ_{Gauss} (with a fixed number of particles for passive FLO). For such states $\operatorname{tr}(\tilde{\mathbb{P}}\Psi_{\text{Gauss}}\otimes\Psi_{\text{Gauss}}) = 1$ and therefore for such we cannot get strong anticoncentration inequalities using Eq. (35). We have also tried to use higher moments in conjugation with the Payley-Zygmund inequality but this did not work. We leave the question whether FLO circuits anticoncentrate when acting on Gaussian states as an open problem.

VI. HARDNESS OF SAMPLING

In this part we use anticoncentratio of FLO circuits and standard complexity-theoretic conjectures to prove classical hardness for sampling from FLO circuits initialized by magic states. We adopt to the fermionic setting standard techniques [12,17,18,20] that use the anticoncentration property to prove hardness of sampling based on conjectures about hardness of approximation of probability amplitudes $p_{\mathbf{x}}(V,\Psi_{\text{in}})$ to within relative error.

We start with a formal definition of a sampling problem defined by FLO circuits initialized in magic input states.

Definition 2: (Fermion sampling task). Let $\mathcal{H}_{\text{pas}} = \bigwedge^{2N}(\mathbb{C}^{4N})$ and let $\mathcal{H}_{\text{act}} = \mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^{4N})$ be Hilbert spaces

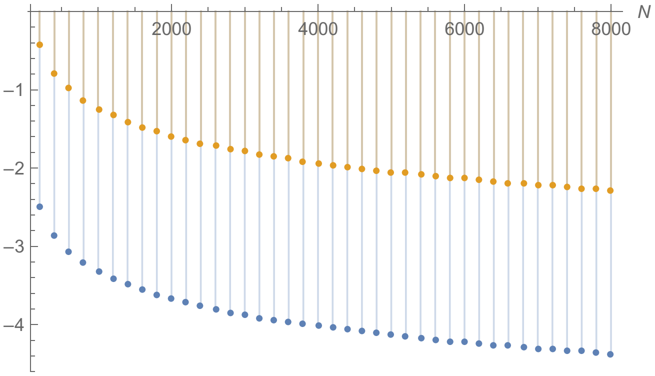


FIG. 7. Plots of the logarithm of the expression (E103) (blue) and $\log(C_{\text{act}}/\sqrt{N}) = \log(16.2/\sqrt{\pi N})$ (orange), which is a valid upper bound for all $N \leq 7000$.

describing $2N$ fermions in $4N$ modes and positive parity fermions in $4N$ modes. Let $\mathcal{G}_{\rm pas}$ and $\mathcal{G}_{\rm act}$ be passive and active FLO transformation. Let V be a FLO circuit on the Hilbert space $\mathcal{H}_{\rm pas}$ or $\mathcal{H}_{\rm act}$ and let $p(V)$ denote probability distribution $p_{\mathbf{x}}(V, \Psi_{\rm in})$. Given a description of V, sample from a probability distribution $q(V)$ that is ϵ close to $p(V, \Psi_{\rm in})$ in l_1 norm (twice the total variation distance)

$$\|p(V) - q(V)\|_1 = \sum_{\mathbf{x}} |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| \le \epsilon, \tag{40}$$

in time poly(N).

Remark 7: It is more convenient to use l_1 norm in place of the total variation distance (TVD) as it appears more directly in the proof of Theorem 2.

It was realized in Refs. [10,12] that, by virtue of Stockmeyer's theorem, the hardness of classically sampling from $p_{\mathbf{x}}(V, \Psi)$ up to an additive error is connected to the hardness of computing $p_{\mathbf{x}}(V, \Psi)$ for most instances of $\mathbf{x}$ and U. In particular, the existence of a classical machine that performs the sampling task implies average-case approximation in a low level of the complexity class called the polynomial hierarchy. To prove this fact, we start by defining the notion of approximating in the average case.

Definition 3: An algorithm $\mathcal{O}$ is said to give an (η, δ)-multiplicative approximate of $q_{\mathbf{z}}$ on average over the probability distribution $\mathcal{P}$ of inputs $\mathbf{z}$ if and only if $\mathcal{O}$ outputs $\mathcal{O}_{\mathbf{z}}$ such that

$$\Pr_{\mathbf{z}\sim\mathcal{P}} [|\mathcal{O}_{\mathbf{z}} - q_{\mathbf{z}}| \le \eta q_{\mathbf{z}}] \ge 1 - \delta. \tag{41}$$

Remark 8: For applications to hardness of sampling, $\mathbf{z}$ will generally be a tuple of inputs $(V, \mathbf{x})$, a FLO circuit and a measurement outcome. Correspondingly, $\mathcal{P}$ will be the joint probability distribution $V \sim \nu_{\rm pas}$ and $\mathbf{x} \sim {\rm unif}(\mathcal{H}_{\rm pas})$ in the case of passive FLO [respectively, $V \sim \nu_{\rm act}$ and $\mathbf{x} \sim {\rm unif}(\mathcal{H}_{\rm act})$ in the case of active FLO], where $\mathbf{x} \sim {\rm unif}(\mathcal{H})$ is the uniform distribution of outcomes restricted to the Hilbert space $\mathcal{H}$.

We now prove the hiding property [10,12,14] of FLO circuits. This allows us to focus on the hardness of a particular outcome probability.

Lemma 1: (Hiding property for FLO.) *Consider a fixed state $|\mathbf{x}_0\rangle \in \mathcal{H}_{\rm pas}$ ($\mathcal{H}_{\rm act}$, respectively) then for any V passive FLO (active FLO, respectively) and $|\mathbf{x}\rangle \in \mathcal{H}_{\rm pas}$ ($\mathcal{H}_{\rm act}$, respectively) there is a passive (active) FLO $V_{\mathbf{x}}$ such that* $|\langle \mathbf{x}| V |\Psi_{\rm in}\rangle|^2 = |\langle \mathbf{x}_0| V_{\mathbf{x}} |\Psi_{\rm in}\rangle|^2$

Proof. It is enough to show that given $\mathbf{x}$ there is $V_{\mathbf{x}}$ passive (active) FLO such that $V_{\mathbf{x}} |\mathbf{x}_0\rangle = |\mathbf{x}\rangle$ up to a global phase. In the passive case this is achieved with gates implementing fermionic swaps $U^{[i,j]}$ such that $U^{[i,j]} f_i^{\dagger} U^{[i,j]\dagger} = f_j^{\dagger}$ and $U^{[i,j]} f_j^{\dagger} U^{[i,j]\dagger} = f_i^{\dagger}$, the order in which they are applied is defined by $|\mathbf{x}\rangle$. The same can be accomplished in the active FLO case with operators $-i m_{2i} m_{2i+1}$ changing the number of fermions (but not parity) and quasi braiding operators $U^{(p,q)}$ to exchange the Majorana operators to the corresponding places. The quasibraidings act on Majorana operators as $U^{(p,q)} m_p (U^{(p,q)})^{\dagger} = m_q$, $U^{(p,q)} m_q (U^{(p,q)})^{\dagger} = m_p$ and $U^{(p,q)} m_x (U^{(p,q)})^{\dagger} = m_x$ when $x \neq p, q$. ■

An additional ingredient required for a quantum sampling advantage is anticoncentration, which states that most output probabilities of a random circuit are sufficiently big so that the approximation error to computing the probabilities is small relative to the probabilities being computed. Both average-case hardness and anticoncentration provide robustness of the sampling task to noise.

Theorem 2: (From approximate sampling to approximately computing probabilities.) *Let $\mathcal{H}_{\rm pas} = \bigwedge^{2N}(\mathbb{C}^{4N})$ and let $\mathcal{H}_{\rm act} = \mathcal{H}^{+}_{\rm Fock}(\mathbb{C}^{4N})$ be Hilbert spaces describing $2N$ Fermions in $4N$ modes and positive parity Fermions in $4N$ modes. Consider in parallel passive FLO circuits and active FLO circuits acting on the input state $|\Psi_{\rm in}\rangle$. If there is a classical algorithm $\mathcal{C}$ that performs fermion sampling as described in Definition 2 with the l_1 error $1/(64C)$, where C is the constant $C_{\rm pas} = 5.7$ (respectively, $C_{\rm act} = 16.2$) appearing in the anticoncentration condition for passive FLO circuits (respectively, active FLO circuits) in Theorem 1.*

Then there is an algorithm in BPP^{NP} (The class BPP^{NP} stands for Probablistic bounded-error classical randomized computation equipped with oracle for solving problems in NP.) that approximates the probability $p_{\mathbf{x}_0}(V, \Psi_{\rm in})$ for an arbitrary but fixed fiducial outcome $\mathbf{x}_0$ up to multiplicative error $1/4 + o(1)$ on $1/(8C)$ fraction of FLO circuits drawn from the distribution $\nu = \nu_{\rm pas}$ for passive FLO circuits (respectively, $\nu_{\rm act}$ for active FLO circuits.)

The proof of the above theorem is given in the Appendix and follows the standard reduction based on Stockmayer algorithm [12,20]. Alternatively, one could arrive at a similar result in two steps: first showing that a classical approximate sampler implies approximations up to an additive error $\epsilon/|\mathcal{H}|$, where ϵ is the TV distance achieved in the sampling task in the polynomial hierarchy, then showing that anticoncentration improves the approximations to multiplicative ones [14]. The alternative proof may be beneficial when anticoncentration does not hold or is undesirable, for example, when anticoncentration renders (black box) certification of quantum advantage infeasible [61].

Armed with Theorem 2, we now state the other conjectures needed before proving the hardness of sampling.

Conjecture 1: (Average case of approximating probabilities on FLO circuits initialized in $|\Psi_{\mathrm{in}}\rangle$). Computing a $[1/4+o(1), 1/(8C)]$-multiplicative approximate to $p_{\mathbf{x}_0}(V, \Psi_{\mathrm{in}})$ for $1/(8C)$ fraction of V sampled from the Haar distribution ν is $\#P$-hard. ($C = C_{\mathrm{pas}}, \nu = \nu_{\mathrm{pas}}$ for passive FLO circuits and $C = C_{\mathrm{act}}, \nu = \nu_{\mathrm{act}}$ for active FLO circuits).

Conjecture 2: The polynomial hierarchy does not collapse.

Remark 9: The motivation for Conjecture 1 comes from the fact that computing exactly the probabilities is $\#P$-hard, this can be seen by writing the output as a polynomial as in Lemma 10, it has been shown that computing a permanent exactly reduces to computing this polynomial and it is known that computing the permanent exactly is $\#P$-hard.

Theorem 3: (Hardness of sampling from FLO circuits initialized in $|\Psi_{\mathrm{in}}\rangle$.) *If Conjectures 1 and 2 are true, then there is no efficient classical algorithm that can approximately sample with l_1 error $1/(64C_{\mathrm{pas}})$ [respectively, $1/(64C_{\mathrm{act}})$] from output probability distributions induced by passive (respectively, active) FLO circuits with the input given by $|\Psi_{\mathrm{in}}\rangle$.*

Proof. By Theorem 2, if there were an approximate sampler with respect to passive (respectively, active) FLO circuits with input $|\Psi_{\mathrm{in}}\rangle$, then there would exist an algorithm BPP^{NP} that $[1/4+o(1), 1/(8C)]$-multiplicative approximates $p_{\mathbf{x}_0}(V, \Psi_{\mathrm{in}})$ in for $1/(8C)$ fraction of passive (respectively, active) FLO circuits. Where $C = C_{\mathrm{pas}}$ in the passive case and $C = C_{\mathrm{act}}$ in the active. By Conjecture 1 this is a $\#P$-hard problem. It is known [104] that BPP is inside the third level of the polynomial hierarchy, i.e., $BPP^{NP} \subseteq \Sigma_3$. By a well-known result of Toda [105] $PH \subseteq P^{\#P}$ and thus $PH \subseteq \Sigma_3$. ■

VII. CAYLEY PATH FOR UNITARY AND ORTHOGONAL GROUPS

In this section, following Ref. [15], we introduce a rational interpolation between elements of the low-dimensional symmetry groups underlying FLO transformations. In what follows by G we denote either of the Lie group $\mathrm{U}(d)$ or $\mathrm{SO}(2d)$. The rational interpolation is constructed from the Cayley transform, which is a rational mapping form the Lie algebra $\mathfrak{g}$ into the corresponding group G. For both groups we give upper bounds for the TVD between the Haar measure μ_G on G and and its *deformations* μ_G^{θ} obtained via Cayley path. These bounds imply TVD bounds between distributions of the corresponding FLO circuits. This and other technical results established below will be called upon in the proof of the worst-to-average-case reduction in Sec. VIII.

The Lie algebras $\mathfrak{u}(d)$ and $\mathfrak{so}(2d)$ of $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ are defined to be

$$\mathfrak{u}(d) = \{X \in \mathbb{C}^{d\times d} | X^{\dagger} = -X\}, \tag{42}$$

$$\mathfrak{so}(2d) = \{X \in \mathbb{R}^{2d\times 2d} | X^{T} = -X\}, \tag{43}$$

where X^T denotes the transpose of the matrix X.

Remark 10: We do not use the physicists' convention, which requires that elements of Lie algebra X satisfy $\exp(i\theta X) \in G$. Therefore, in particular, here $\mathfrak{u}(d)$ [respectively, $\mathfrak{so}(d)$] consists of skew-Hermitian (respectively, antisymmetric) matrices.

Every element $X \in \mathfrak{g}$ defines a one-parameter path in G: $\{\exp(\theta X)\}_{\theta\in\mathbb{R}}$, via the exponential map, $\exp : \mathfrak{g} \to G$. Both orthogonal and unitary groups are compact and connected. Therefore, exponential map is surjective and can be used to parametrize G, and provides an interpolation between any two group elements. However, the interpolation is not polynomial in nature, and while it is possible to truncate the power series of exp to obtain a polynomial interpolation [14], the resulting interpolation represents circuits that are not unitary (cf. [15]).

To remedy this, Ref. [15] employs an algebraic Cayley transformation between $\mathfrak{u}(d)$ and $\mathrm{U}(d)$. This transformation can be however defined more generally as a mapping between Lie algebra and the corresponding Lie group [106]. For our needs it is enough to consider the case of unitary and special orthogonal groups.

Definition 4: Let G be $\mathrm{U}(d)$ or $\mathrm{SO}(2d)$, and let $\mathfrak{g}$ denotes its Lie algebra. The Cayley transform is a mapping $f : \mathfrak{g} \to G$ defined via

$$f(X) = (\mathbb{I} - X)(\mathbb{I} + X)^{-1}. \tag{44}$$

It is easy to see that the image of $f(\mathfrak{g})$ equals a dense subset $\tilde{G} = \{g \in G \,|\, \{-1\} \notin sp(g)\}$ consisting of elements of G (i.e., unitary or orthogonal matrices) that do not have -1 in their spectrum. On $\tilde{G}$ the inverse of f is well defined. Specifically, $f^{-1} : \tilde{G} \to \mathfrak{g}$ is given by

$$f^{-1}(g) = (\mathbb{I} - g)(\mathbb{I} + g)^{-1}, \tag{45}$$

where $g \in \tilde{G}$. This explicit form of the inverse map can be verified directly from the definition of f. Cayley map defines a path deformation between $g_0 \in G$ and $g_0 f(X)$ as follows (see Fig. 8).

Cayley map can be used to define a rational interpolation between arbitrary group elements. To this end consider first the map $F_\theta : \tilde{G} \to G$, given by

$$F_\theta(g) = f[\theta f^{-1}(g)], \; \theta \in [0, 1]. \tag{46}$$

The above mapping can be evaluated explicitly (note that elements of the considered Lie groups are normal matrices

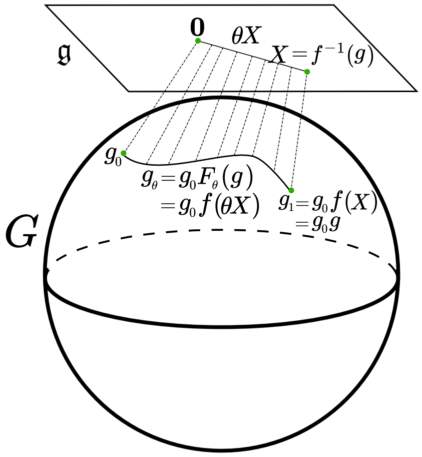


FIG. 8. Path deformation defined by the Cayley map in Eq. (44). A path is induced between element $g_0 \in G$ and g_0g by taking $X = f^{-1}(g) \in \mathfrak{g}$ and considering the perturbation $g_\theta = g_0f(\theta X)$.

and therefore functional calculus can be performed effectively in the same way as if we were dealing with functions of a real variable):

$$F_\theta(g) = \frac{(1-\theta)\mathbb{I} + (1+\theta)g}{(1+\theta)\mathbb{I} + (1-\theta)g},\ \theta \in [0,1]. \tag{47}$$

For both orthogonal and unitary operators we have $\|g\| = 1$. Therefore, for $\theta \in (0,1]$ the denominator of (47) does not vanish and therefore we can use (47) to define F_θ to be a function defined on whole G, while for any $g \in G$ we get that $\lim_{\theta\to 0} F_\theta(g) = \mathbb{I}$. Therefore, for $\theta \in [0,1]$ the denominator of Eq. (47) does not vanish and therefore we can use Eq. (47) to define F_θ as a function defined on whole G. Importantly, for the fixed input as θ goes from 0 to 1 we move on a rational path form the identity $\mathbb{I}$ to g. Consequently the path

$$g_\theta = g_0F_\theta(g),\ \theta \in [0,1] \tag{48}$$

is a rational interpolation between a fixed group element g_0 (which can correspond, for example, to a worst-case #P-hard FLO circuit) and a completely generic group element g_0g.

It is important to note that both $f(\theta X)$ and X can be simultaneously brought into a block-diagonal form by conjugation by elements of the group: $M \mapsto gMg^{-1}$. It follows from the fact that $f(\theta X)$ is simply a function of X and then the transformation properties of elements of $\mathfrak{g}$ under the conjugation by elements of G. For the case of $X \in \mathfrak{u}(d)$ we have an elementary fact from linear algebra that there exist a unitary $U \in \mathrm{U}(d)$ such that

$$UXU^\dagger = \sum_{j=1}^{d} \phi_j X_j, \tag{49}$$

where $X_j = i|j\rangle\langle j|$. Similarly, for any $X \in \mathfrak{so}(2d)$ there exist $O \in \mathrm{SO}(2d)$ such that

$$OXO^T = \sum_{j=1}^{d} \phi_j \tilde{X}_j, \tag{50}$$

where $\tilde{X}_j = |2j\rangle\langle 2j-1| - |2j-1\rangle\langle 2j|$ is the generator for the jth block. These statements have analogs on the level of elements of the group. Every unitary U can be transformed into a diagonal form

$$\mathrm{diag}(e^{i\phi_1}, e^{i\phi_2}, \ldots, e^{i\phi_d}) = \exp\left(\sum_{j=1}^{d} \phi_j X_j\right). \tag{51}$$

For elements $\mathrm{SO}(2d)$, the block diagonalization amounts to the geometric fact that any $2d$-dimensional rotation can be decomposed into d-independent planar rotations of the form $\exp\left(\sum_{j=1}^{d} \phi_j \tilde{X}_j\right)$.

The following lemma, which we prove in Appendix C, shows that for $1-\theta \le o(1/d^2)$ the distribution of elements of the group g, and $g_\theta = g_0F_\theta(g)$, where $g \sim \mu_G$, are close in total variation distance.

Lemma 2: (TV distance between the Haar measure in G and its θdeformation.) *Let G be equal to $U(d)$ or $SO(2d)$. Let $g_0 \in G$ be a fixed element in G. Let $g \sim \mu_G$ an let $g_\theta = g_0F_\theta(g)$, for $\theta \in [0,1]$ and $F_\theta : G \to G$ defined in Eq. (47). Let now μ_G^θ denotes the induced measure according to which g_θ is distributed. Assume furthermore that $\theta \in [1-\Delta, 1]$, for $\Delta > 0$. We then have*

$$\begin{aligned} \left\|\mu_{U(d)} - \mu_{U(d)}^\theta\right\|_{\mathrm{TVD}} &\le d^2\Delta/2, \\ \left\|\mu_{SO(2d)} - \mu_{SO(2d)}^\theta\right\|_{\mathrm{TVD}} &\le d^2\Delta/2. \end{aligned} \tag{52}$$

Remark 11: A similar analysis was carried out in Ref. [15] for the case of unitary group $\mathrm{U}(d)$. There, however, considerations were carried out for $d = O(1)$. This was justified because gates in question were only single- and two-qubit gates. The above lemma can be viewed as an extension of the analysis given there in the sense of allowing arbitrary relation between d and Δ.

The robustness of the quantum supremacy claim will be tied directly to the degree of the rational functions that interpolate between quantum circuits (Appendix D). Here we give the explicit rational functions and their degrees in the Cayley-path interpolation $g_\theta = F_\theta(g)$ at the group level, Eq. (47), in U(d) and SO($2d$). [A similar result for U(d) was derived in Ref. [15]].

In the case of U(d), since g can always be diagonalized by some element h: $hgh^{-1} = \sum_{j=1}^{d} e^{i\phi_j} \, |j\rangle\langle j|$, we have that

$$g_\theta = \sum_{j=1}^{d} \frac{(1-\theta)+(1+\theta)e^{i\phi_j}}{(1+\theta)+(1-\theta)e^{i\phi_j}} g_0 h^{-1} \, |j\rangle\langle j| \, h \tag{53}$$

$$= \sum_{j=1}^{d} \frac{1+i\theta\tan(\phi_j/2)}{1-i\theta\tan(\phi_j/2)} g_0 h^{-1} \, |j\rangle\langle j| \, h \tag{54}$$

$$= \frac{1}{\mathcal{Q}_g(\theta)} \sum_{j=1}^{d} P_j(\theta) g_0 h^{-1} \, |j\rangle\langle j| \, h =: \frac{\mathcal{P}_{g_0,g}(\theta)}{\mathcal{Q}_g(\theta)}, \tag{55}$$

where

$$\mathcal{Q}_g(\theta) = \prod_{j=1}^{d} [1 - i\theta\tan(\phi_j/2)], \tag{56}$$

$$P_j(\theta) = [1 + i\theta\tan(\phi_j/2)] \prod_{\substack{1\le k\le d \\ k\neq j}} [1 - i\theta\tan(\phi_k/2)] \tag{57}$$

are both polynomials of degree d in θ, and $\mathcal{P}_{g_0,g}(\theta)$ is a formal polynomial that depends on the matrices g and g_0.

The same calculation applies to the case SO($2d$), except that now each eigenspace is two dimensional and spanned by $\mathbb{I}_j = |2j-1\rangle\langle 2j-1| + |2j\rangle\langle 2j|$ and $\tilde{X}_j = |2j\rangle\langle 2j-1| - |2j-1\rangle\langle 2j|$. Again, let $hgh^{-1} = \sum_{j=1}^{d}(\cos\phi_j \mathbb{I}_j + \sin\phi_j \tilde{X}_j)$, and one has

$$\begin{aligned} g_\theta = g_0 \sum_{j=1}^{d} &[1+\cos\phi_j + \theta^2(1-\cos\phi_j)]^{-1} \\ &\times \left\{\left[1+\cos\phi_j - \theta^2(1-\cos\phi_j)\right]\mathbb{I}_j \right. \\ &\left. + 2\theta \sin\phi_j h^{-1}\tilde{X}_j h\right\} \end{aligned} \tag{58}$$

$$= g_0 \sum_{j=1}^{d} \frac{[1-\theta^2\tan^2(\phi_j/2)]\mathbb{I}_j + 2\theta\tan(\phi/2)h^{-1}\tilde{X}_j h}{1+\theta^2\tan^2(\phi_j/2)} \tag{59}$$

$$= \frac{1}{\mathcal{Q}_g(\theta)} \sum_{j=1}^{d} \left(P_j^{\mathrm{diag}}(\theta) g_0 \mathbb{I}_j + P_j^{\mathrm{off}}(\theta) g_0 h^{-1}\tilde{X}_j h\right) \tag{60}$$

$$=: \frac{\mathcal{P}_{g_0,g}(\theta)}{\mathcal{Q}_g(\theta)}, \tag{61}$$

where in Eq. (59) we divided both the numerator and the denominator by $1+\cos\phi_j$ and

$$\mathcal{Q}_g(\theta) = \prod_{j=1}^{d} [1+\theta^2\tan^2(\phi_j/2)], \tag{62}$$

$$P_j^{\mathrm{diag}}(\theta) = [1+\theta^2\tan^2(\phi_j/2)]^2 \times \prod_{\substack{1\le k\le d \\ k\neq j}} [1+\theta^2\tan^2(\phi_j/2)]^2, \tag{63}$$

$$P_j^{\mathrm{off}}(\theta) = 2\theta\tan(\phi_j/2) \prod_{\substack{1\le k\le d \\ k\neq j}} [1+\theta^2\tan^2(\phi_j/2)]^2 \tag{64}$$

are polynomials in θ of degree $2d$, $2d$, and $2d-1$, respectively, and $\mathcal{P}_{g_0,g}(\theta)$ is a formal polynomial that depends on the matrices g and g_0.

Below we give a lower bound for $\mathcal{Q}_g(\theta)$ to assure that the rational function does not blow up, and an upper bound for generic $g \in G$, which will be crucial for a robust reduction in Sec. VIII. Note that the coefficients of the polynomial $\mathcal{Q}_g(\theta)$ depends only on generalized eigenvalues of g ($e^{i\phi_j}$ in the unitary case and $\cos\phi_j$, $\sin\phi_j$ in the orthogonal case) and hence $Q(\theta)$ can be precomputed in time polynomial in d by diagonalizing g, computing each $\tan(\phi_j/2)$, which is just an algebraic function of $e^{i\phi_j}$, and computing the final result.

Lemma 3: *Let $\mathcal{Q}_g(\theta)$ be the polynomial in defined in Eq. (56) for $G = U(d)$ and in Eq. (62) for $G = SO(2d)$. Let now $\tilde{\Delta} > 0$. Then we have the following inequalities:*

$$\Pr_{g\sim\mu_{U(d)}} \left\{ \left|\mathcal{Q}_g(\theta)\right|^2 \le \left[1+\left(\frac{\theta\pi}{\tilde{\Delta}}\right)^2\right]^d \right\} \ge 1 - d\frac{\tilde{\Delta}}{\pi}, \tag{65}$$

$$\Pr_{g\sim\mu_{SO(2d)}} \left\{ \left|\mathcal{Q}_g(\theta)\right|^2 \le \left[1+\left(\frac{\theta\pi}{\tilde{\Delta}}\right)^2\right]^{2d} \right\} \ge 1 - d\frac{\tilde{\Delta}}{\pi}. \tag{66}$$

In addition, for all g, $|\mathcal{Q}_g(\theta)|^2 \ge 1$ for both $U(d)$ and $SO(2d)$.

Proof. Since $g \in G$ is Haar distributed, every generalized eigenphase ϕ_j is distributed uniformly on the interval $[-\pi,\pi]$ [107]. Therefore, for every j we have

$$\Pr_{g\sim\mu_G}\left(\phi_j \in [\pi-\tilde{\Delta},\pi] \cup [-\pi,-\pi+\tilde{\Delta}]\right) = \frac{\tilde{\Delta}}{\pi}. \tag{67}$$

It is easy to verify that for $\phi_j \in [-\pi+\tilde{\Delta}, \pi-\tilde{\Delta}]$ we have $\left|\tan(\phi_j/2)\right| \le \pi/\tilde{\Delta}$. Using the union bound over different ϕ_j, $j \in [d]$ we obtain that with probability at least

$1 - d(\tilde{\Delta}/\pi)$,

$$\left|\tan(\phi_j/2)\right| \leq \pi/\tilde{\Delta} \text{ for all } j \in [d]. \tag{68}$$

Using the definition of polynomials $\mathcal{Q}_g(\theta)$ from Eqs. (56) and (62), we obtain the claimed inequalities (65) and (66). ■

Remark 12: We believe that the inequalities stated in Lemma 3 can be greatly improved by the usage of more sophisticated techniques from random matrix theory. However, for our purposes these crude estimates are sufficient (see proof of Theorem 7).

VIII. ROBUST AVERAGE-CASE HARDNESS OF OUTPUT PROBABILITIES OF FERMIONIC CIRCUITS

In this part we give strong evidence for Conjecture 1 used to show classical hardness of sampling from fermionic linear circuits initialized in $|\Psi_{\text{in}}\rangle$ (cf. Theorem 3). There we conjectured that it is #P-hard to approximate probabilities $p_{\mathbf{x}}(V,\Psi_{\text{in}}) = |\langle\mathbf{x}|\,V|\Psi_{\text{in}}\rangle|^2$ of *generic* FLO circuits initialized in $|\Psi_{\text{in}}\rangle$ to relative error. To support the conjecture we prove weaker theorems showing average-case #P-hardness of exact computation of $p_{\mathbf{x}}(V,\Psi_{\text{in}})$ (Theorem 5) and extend it further to average-case #P-hardness of approximating $p(\mathbf{x}|V,\Psi_{\text{in}})$ up to error $\epsilon = \exp[-\Theta(N^6)]$ (Theorem 7), where N is the number of states $|\Psi_4\rangle$ used.

To establish this we combine previously known worst-case hardness results (see discussion in Sec. VI) and adopt to our setting rational interpolation method based on Cayley transform introduced recently by Movassagh [15]. We also use the fact that both passive FLO circuits ($\mathcal{G}_{\text{pas}}$) as well active FLO ($\mathcal{G}_{\text{act}}$) are representations of low-dimensional symmetry group G [equal to U(d) or SO($2d$)]. As shown in the previous section, this implies that, when evaluated on the Cayley path g_θ, the circuit rise to outcome probabilities being rational functions of low degree (cf. Lemma 11). This low-degree structure allows worst-to-average-case reductions to be performed for the family of circuits considered. Specifically, we reduce the problem of computation of the worst-case probability $p(\mathbf{x}|V_0,\Psi_{\text{in}})$ to computing $p(\mathbf{x}|V,\Psi_{\text{in}})$, for V being typical passive or active FLO circuit.

We need the following result that guarantees that it is possible to recover an unknown rational function $F(\theta)$ from a set of its values at different points, even if some of the evaluation are erroneous.

Theorem 4: (Berlekamp-Welch for rational functions [15].) *Let $R(\theta)$ be a rational function of degree* $\deg(R) = (d_1, d_2)$. *A set of points $\mathcal{S} = \{(\theta_1,r_1),(\theta_2,r_2),\ldots,(\theta_L,r_L)\}$ specifies $R(\theta)$ uniquely provided $L > d_1 + d_2 + 2t$, where*

$$|\{i \in [L]\,|\,R(\theta_i) \neq r_i\}| \leq t. \tag{69}$$

Moreover, $R(\theta)$ can be recovered in polynomial time in L and $\deg(R)$, *when $\mathcal{S}$ is given.*

Recall that in our quantum advantage scheme we have $d = 4N$, $|\Psi_{\text{in}}\rangle = |\Psi_4\rangle^{\otimes N}$, for $|\Psi_4\rangle = (|0011\rangle + |1100\rangle)/\sqrt{2}$ (therefore for the case of passive FLO $n = 2N$). Let now $g_0 \in G$ be an element of the symmetry group such that $p_{\mathbf{x}_0}(V_0,\Psi_{\text{in}})$ is #P-hard to compute, where $V_0 = \Pi(g_0)$ and $\mathbf{x}_0$ is the specific output state. We use a Cayley-path interpolation between g_0 and Haar-random elements from G

$$g_\theta = g_0 F_\theta(g),\ g \sim \mu_G. \tag{70}$$

Let μ_G^θ be the distribution of g_θ obtained in this way. In Lemma 2 we proved bounds for the TV distances $\|\mu_G - \mu_G^\theta\|_{TVD}$. These bounds can be directly translated on the level of the corresponding circuits. Indeed, let $V_\theta = \Pi(g_\theta)$ and let ν_G^θ denote the distribution of the corresponding quantum circuits obtained by appropriate representation Π of G. Since distribution of the Haar random FLO circuits $\nu_{\text{pas}}, \nu_{\text{pas}}$ are obtained in exactly the same way we get from the monotonicity of TV distance (cf. Sec. II).

$$\begin{aligned} \left\|\nu_{\text{pas}} - \nu_{\text{pas}}^\theta\right\|_{\text{TVD}} &\leq 8N^2\Delta, \\ \left\|\nu_{\text{act}} - \nu_{\text{act}}^\theta\right\|_{\text{TVD}} &\leq 8N^2\Delta, \end{aligned} \tag{71}$$

where $\theta \in [1-\Delta, 1]$. Finally, from Lemma 11 we know that probabilities $R(\theta) = \text{tr}[|\mathbf{x}_0\rangle\langle\mathbf{x}_0|\Pi(g_\theta)\rho\Pi(g_\theta)^\dagger]$ are rational functions of the deformation parameter θ of degrees,

$$\begin{aligned} \text{passive FLO:}\quad \deg(R) &= (16N^2, 16N^2), \\ \text{active FLO:}\quad \deg(R) &= (32N^2, 32N^2), \end{aligned} \tag{72}$$

and act that passive and active fermionic linear circuits give an average to worst-case reduction for outcome probabilities generated by fermionic circuits.

Theorem 5: (Average-case #P-hardness of computation of outcome probabilities of FLO circuits.) *Let V_0 be a FLO circuit such that computing $p_{\mathbf{x}_0}(V_0,\Psi_{\text{in}}) = |\langle\mathbf{x}_0|\,V_0\,|\Psi_{\text{in}}\rangle|^2$ is #P-hard, where V_0 is element of either passive or active FLO circuits and the output Fock state $|\mathbf{x}_0\rangle$ belongs to the suitable Hilbert space $\mathcal{H}_{\text{pas}} = \bigwedge^{2N}(\mathbb{C}^{4N})$ for passive FLO and $\mathcal{H}_{\text{act}} = \mathcal{H}^+_{\text{Fock}}(\mathbb{C}^{4N})$, respectively.*

Then it is #P-hard to compute $p_{\mathbf{x}_0}(V,\Psi_{\text{in}}) = |\langle\mathbf{x}_0|\,V\,|\Psi_{\text{in}}\rangle|^2$ with probability $\alpha > 3/4 + \delta$, $\delta = 1/\text{poly}(N)$, over the the uniform distribution of circuits: $V \sim \nu_{\text{pas}}$ for passive FLO and $V \sim \nu_{\text{act}}$ for active FLO.

Remark 13: Due to hiding property (see Lemma 1) both active and passive FLO gates can permute between possible output Fock states $|\mathbf{x}\rangle$ in $\mathcal{H}_{\text{pas}} = \bigwedge^{2N}(\mathbb{C}^{4N})$ and $\mathcal{H}_{\text{act}} = \mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^{4N})$, respectively. Therefore, using the invariance of the Haar measure on G, we can transform $\mathbf{x}_0$ above into any other output $\mathbf{x}$ satisfying $|\mathbf{x}| = 2N$ (for passive FLO) and $|\mathbf{x}|$ even (for active FLO).

Proof. We first fix the symmetry group G describing a class of FLO circuits. The proof is virtually identical for both $G = \text{U}(4N)$ and $G = \text{SO}(8N)$. Suppose that $\mathcal{O}$ is an oracle that given a description of $\Pi(g)$ computes $|\langle\mathbf{x}_0|\,\Pi(g)\,|\Psi_{\text{in}}\rangle|^2$ with high probability, i.e.,

$$\Pr_{g\sim\mu_G}\left[\mathcal{O}[\Pi(g)] = |\langle\mathbf{x}_0|\,\Pi(g)\,|\Psi_{\text{in}}\rangle|^2\right] > \alpha. \tag{73}$$

The uniform distribution of FLO circuits ν_G is obtained by setting $V = \Pi(g)$, where $g \sim \mu_G$ [recall that $\nu_G = \nu_{\text{pas}}$ for $G = \text{U}(4N)$ and $\nu_G = \nu_{\text{act}}$ for $G = \text{SO}(8N)$]. Therefore, Eq. (73) is equivalent to

$$\Pr_{V\sim\nu_G}\left[\mathcal{O}(V) = |\langle\mathbf{x}_0|\,V|\Psi_{\text{in}}\rangle|^2\right] > \alpha. \tag{74}$$

In what follows we argue that oracle $\mathcal{O}$ can be used to compute the #P-hard probability in polynomial time. The argument presented below follows steps from worst-to-average-case reduction for permanents of Gaussian matrices from Ref. [10], and its modification that involving rational interpolation from Ref. [15]. We consider a rational path interpolation $g_\theta = g_0 F_\theta(g)$ between worst case g_0 and $=g' = g_0 g$, where g is chosen according to Haar measure on G. We call $\mathcal{O}$ on L distinct FLO circuits $\Pi(g_{\theta_1}), \Pi(g_{\theta_2}), \ldots, \Pi(g_{\theta_L})$, where $\theta_i \in [1-\Delta, 1]$, and the parameter Δ will be chosen later. We use evaluations $\{\mathcal{O}[\Pi(g_{\theta_i})]\}_{i=1}^{L}$ to efficiently reconstruct the rational function using Berlekamp-Welch algorithm for rational functions $R(\theta) = |\langle\mathbf{x}_0|\,\Pi(g_\theta)\,|\Psi_{\text{in}}\rangle|^2$ (cf. Theorem 4). If the reconstruction is successful, evaluation of $R(\theta)$ at $\theta = 0$ gives us the #P-hard probability $R(0) = |\langle\mathbf{x}_0|\,V_0\,|\Psi_{\text{in}}\rangle|^2$ [we use here $V_0 = \Pi(g_0)$].

To assess the success probability with which the above scheme evaluates $|\langle\mathbf{x}_0|\,V_0\,|\Psi_{\text{in}}\rangle|^2$ correctly we first bound the success probability with which oracle $\mathcal{O}$ computes the value of $p_{\mathbf{x}_0}[\Pi(g_\theta), \Psi_{\text{in}}]$ correctly. Using variational characterization of TV distance and bounds from Eq. (71) we obtain

$$\Pr_{V\sim\nu_G}\left[\mathcal{O}(V) = |\langle\mathbf{x}_0|\,V|\Psi_{\text{in}}\rangle|^2\right] - \Pr_{V\sim\nu_G^{\theta}}\left[\mathcal{O}(V) = |\langle\mathbf{x}_0|\,V|\Psi_{\text{in}}\rangle|^2\right] \leq CN^2\Delta. \tag{75}$$

Combining the above with Eq. (74) we get

$$\Pr_{V\sim\nu_G^{\theta}}\left[\mathcal{O}(V) = |\langle\mathbf{x}_0|\,V|\Psi_{\text{in}}\rangle|^2\right] \geq \alpha - CN^2\Delta. \tag{76}$$

Or equivalently

$$\Pr_{g\sim\mu_G}\left[\mathcal{O}[\Pi(g_\theta)] = |\langle\mathbf{x}_0|\,\Pi(g_\theta)\,|\Psi_{\text{in}}\rangle|^2\right] \geq \alpha - CN^2\Delta. \tag{77}$$

According to rational Berlekamp-Welch algorithm the number of evaluations L of a rational function $R(\theta)$ that allows reconstruction of it despite having at most t incorrect evaluations has to satisfy $L > d_1 + d_2 + 2t$. Note that in the considered case $d_1 + d_2 = \Theta(N^2)$ [cf. Eq. (72)]. The probability of having a number of errors that exceeds the bound allowing for reconstruction of $R(\theta)$ can be estimated using Markov inequality applied for the random variable counting the number of invalid evaluations of the oracle

$$t(g) = \left|\left\{\theta_i \mid \mathcal{O}[\Pi(g_{\theta_i})] \neq |\,\langle\mathbf{x}_0|\,\Pi(g_{\theta_i})\,|\Psi_{\text{in}}\rangle\,|^2,\ i \in [L]\right\}\right|. \tag{78}$$

From the definition of t and the inequality (77) it follows that $\mathbb{E}_{g\sim\mu_G} t(g) \leq [1 - \alpha + CN^2\Delta]L$. Using this estimate in Markov inequality [recall that by assumption $\alpha > 3/4 + \delta$, for $\delta = 1/\text{poly}(N)$] we get

$$\Pr_{g\sim\mu_G}\left[t(g) > \frac{L - d_1 - d_2}{2}\right] \leq \frac{[1 - \alpha + CN^2\Delta]L}{\frac{L-d_1-d_2}{2}} \leq \frac{\frac{1}{4} - \delta + CN^2\Delta}{\frac{1}{2} - \frac{d_1+d_2}{2L}}. \tag{79}$$

By choosing Δ and L such that $CN^2\Delta \leq \delta/2$ and $d_1 + d_2/2L \leq \delta/4$ [this can be done with $\Delta = 1/\text{poly}(N)$ and $L = \text{poly}(N)$ because $d_1 + d_2 = \Theta(N^2)$], we obtain

$$\Pr_{g\sim\mu_G}\left[t(g) > \frac{L - d_1 - d_2}{2}\right] \leq \frac{\frac{1}{4} - \frac{\delta}{2}}{\frac{1}{2} - \frac{\delta}{4}} \leq \frac{1}{2} - \frac{\delta}{4}. \tag{80}$$

The leftmost part of the above inequality is the probability of failure of our protocol. Therefore, since $\delta = 1/\text{poly}(N)$, we can repeat the procedure polynomially numerous times, for different choices of $\Pi(g)$, compute $R\Pi(g)(0)$ each time, and output the majority vote. The probability of successfully computing the right result (i.e., $|\langle\mathbf{x}_0|\,V_0\,|\Psi_{\text{in}}\rangle|^2$) can be made exponentially close to 1 in this way. ■

We proceed with proving the robust version of the above result. To this end, we shift to polynomial interpolation because much more is known about its robustness to errors. To phrase our problem using polynomials we first note that the rational function $R_{g_0,g} = |\langle\mathbf{x}_0|\,\Pi(g_\theta)\,|\Psi_{\text{in}}\rangle|^2$, where $g_\theta = g_0 F_\theta(g)$ can be written as

$$R_{g_0,g}(\theta) = \frac{D_{g_0,g}(\theta)}{Q_g(\theta)}, \tag{81}$$

where for both groups $D_{g_0,g}, Q_g$ are real polynomials of degrees $D_{g_0,g} = d_1 = \Theta(N^2)$, $Q_g = d_2 = \Theta(N^2)$ (cf.

Lemma 11). Moreover, the denominator $Q_g(\theta)$ can be computed efficiently in N [see Eq. (D22)] given a classical description of g. Hence, we have the following.

Lemma 4: *Let $R_{g_0,g}(\theta)$ be defined as in Eq. (81) [for fixed $g_0, g \in G$, where $G = U(4N)$ or $G = SO(8N)$]. Then complexity of computation of $R_{g_0,g}(\theta)$ and $D_{g_0,g}(\theta)$ is equivalent up to $\Theta(N^2)$ overhead.*

The above allows us to use, following Refs. [14,15], techniques of polynomial interpolation in order to estimate the hard probability $R_{g_0,g}(0)$. We now state two results from this domain that are used later in the robust version of the worst-to-average-case reduction given above.

Lemma 5: (Paturi lemma [108].) *Let $P(\theta)$ be a polynomial of degree k and suppose that $|P(\theta)| \leq \epsilon$ for $\theta \in [1-\Delta, 1]$, $\Delta \in (0,1]$. Then*

$$P(0) \leq \epsilon \exp[4k(1+\Delta^{-1})]. \tag{82}$$

Remark 14: The above lemma is usually presented in a slightly different form in which the assumption $|P(\theta)| \leq \epsilon$ for $\theta \in [-\Delta, \Delta]$ ($\Delta > 0$) is used to establish $P(0) \leq \epsilon \exp[2k(1+\Delta^{-1})]$. Our result can be deduced from the former via simple affine change of variables $\theta \mapsto \theta' = -2/2 - \Delta\theta + 1$.

Theorem 6: (Values of polynomials bounded at equally spaced points [109].) *Let θ_i, $i = 1, \ldots, L$ be a collection of L equally spaced points in the interval $[1-\Delta, 1]$, $\Delta \in (0,1)$. Let $P(\theta)$ be a polynomial of degree k. Assume that for every i, $|P(\theta_i)| \leq \epsilon$. Then there exist absolute constants $a, b > 0$ such that*

$$\max_{\theta \in [1-\Delta, 1]} |P(\theta)| \leq \epsilon \exp\left(b\frac{k^2}{L} + a\right). \tag{83}$$

Remark 15: The problem of bounding values of polynomials that are bounded on a uniformly spaced interval has a long history and there have been more recent developments in this topic (see, for example, Ref. [110]). However, for our purposes the above result by Coppersmith and Rivlin is sufficient.

Before we formulate our result and prove our main theorem we need one more technical ingredient. Informally speaking, since $Q_g(\theta)$ appears in the denominator of Eq. (81) we need to ensure that values of $Q_g(\theta)$ are not too large for typical values of g. This is achieved by combining Lemma 3 and explicit formulas for $Q_g(\theta)$ given in Eq. (D22) we obtain the following.

Corollary 1: *Let $g \in G$ and let $Q_g(\theta)$ be the polynomial in defined in Eq. (D22) for $G = U(d)$ in $G = SO(2d)$. Assume that $n = 2N$, $d = 4N$. Let now $\tilde{\Delta} > 0$. We then have the following inequalities:*

$$\Pr_{g\sim\mu_{U(d)}} \left\{ Q_g(\theta) \leq \left[1 + \left(\frac{\theta\pi}{\tilde{\Delta}}\right)^2\right]^{16N^2} \right\} \geq 1 - 4N\frac{\tilde{\Delta}}{\pi}, \tag{84}$$

$$\Pr_{g\sim\mu_{SO(2d)}} \left\{ Q_g(\theta) \leq \left[1 + \left(\frac{\theta\pi}{\tilde{\Delta}}\right)^2\right]^{32N^2} \right\} \geq 1 - 4N\frac{\tilde{\Delta}}{\pi}. \tag{85}$$

Combining all technical ingredients stated above we are in the position to prove our main result.

Theorem 7: (Average-case #P-hardness of approximation outcome probabilities of FLO circuits.) *Let V_0 be a FLO circuit such that computing $p_{\mathbf{x}_0}(V_0, \Psi_{\text{in}}) = |\langle \mathbf{x}_0 | V_0 | \Psi_{\text{in}} \rangle|^2$ is #P-hard, where V_0 is an element of either passive or active FLO circuits and the output Fock state $|\mathbf{x}_0\rangle$ belongs to the suitable Hilbert space $\mathcal{H}_{\text{pas}} = \bigwedge^{2N}(\mathbb{C}^{4N})$ for passive FLO and $\mathcal{H}_{\text{act}} = \mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^{4N})$, respectively.*

Let $\epsilon = \exp[-\Theta(N^6)]$. Then it is #P-hard to compute $p_{\mathbf{x}_0}(V, \Psi_{\text{in}}) = |\langle \mathbf{x}_0 | V | \Psi_{\text{in}} \rangle|^2$ to accuracy ϵ with probability $\alpha > 1 - \delta$, $\delta = o(N^{-2})$, over the the uniform distribution of circuits: $V \sim \nu_{\text{pas}}$ for passive FLO and $V \sim \nu_{\text{act}}$ for active FLO.

Remark 16: Using the same arguments as in the remark below Theorem 5 we can transform $\mathbf{x}_0$ above into any other output $\mathbf{x}$ satisfying $|\mathbf{x}| = 2N$ (for passive FLO) and $|\mathbf{x}|$ even (for active FLO).

Proof. We first fix the symmetry group G describing a class of FLO circuits. The uniform distribution of FLO circuits ν_G is obtained by setting $V = \Pi(g)$, where $g \sim \mu_G$ [recall that $\nu_G = \nu_{\text{pas}}$ for $G = \text{U}(4N)$ and $\nu_G = \nu_{\text{act}}$ for $G = \text{SO}(8N)$]. The general idea of the proof is similar to the reasoning used to prove Theorem 5. We start with an oracle $\mathcal{O}$ that given a classical description of $V = \Pi(g)$, is able to approximately compute $p_{\mathbf{x}_0}(V, \Psi_{\text{in}}) = |\langle \mathbf{x}_0 | \Pi(g) | \Psi_{\text{in}} \rangle|^2$,

$$\Pr_{g\sim\mu_G} \left[\left| \mathcal{O}[\Pi(g)] - |\langle \mathbf{x}_0 | \Pi(g) | \Psi_{\text{in}} \rangle|^2 \right| \leq \epsilon \right] > 1 - \delta. \tag{86}$$

Equivalently, we have

$$\Pr_{V\sim\nu_G} \left[\left| \mathcal{O}(V) - |\langle \mathbf{x}_0 | V | \Psi_{\text{in}} \rangle|^2 \right| \leq \epsilon \right] > 1 - \delta. \tag{87}$$

For a generic Haar random $g \in G$ we again consider a rational path $g_\theta = g_0 F_\theta(g)$ between $g_0 g$ and g_0, where g_0 is an element of the group corresponding to the worst-case circuit. Recall that by μ_G^θ we denoted the distribution of g_θ for $g \sim \mu_G$. We now query oracle $\mathcal{O}$ multiple times on g_{θ_i}, where θ_i are L equally distributed points in the interval $[1-\Delta, 1]$, for $\Delta > 0$ to be set latter. As done previously in the proof of Theorem 5 by using variational characterization of TV distance and Eq. (71), we obtain that for every $\theta_i \in [1-\Delta, 1]$

$$\Pr_{g\sim\mu_G}\left[\left|\mathcal{O}[\Pi(g_{\theta_i})] - |\langle \mathbf{x}_0|\,\Pi(g_{\theta_i})\,|\Psi_{\rm in}\rangle\,|^2\right| \le \epsilon\right] > 1-\delta-8\Delta N^2, \tag{88}$$

Let now $D_{g_0,g}(\theta)$ be a polynomial of degree $\deg(D_{g_0,g}) = \Theta(N^2)$ that we defined Eq. (81). Recall that the denominator of $R_{g_0,g}(\theta)$, $Q_g(\theta)$ can be computed efficiently (cf. Lemma 4). Therefore, we can use $\mathcal{O}$ to construct an oracle $\tilde{\mathcal{O}}$ that computes approximations of values of polynomial $D_{g_0,g}$ at point θ_i with potentially high probability over the choice of g

$$\Pr_{g\sim\mu_G}\left[|\tilde{\mathcal{O}}[\Pi(g_{\theta_i})] - D_{g_0,g}(\theta_i)| \le \epsilon Q_g(\theta_i)\right] > 1-\delta-8\Delta N^2, \tag{89}$$

We now use Corollary 1 to bound $Q_g(\theta)$ in the above expression:

$$\Pr_{g\sim\mu_G}\left[Q_g(\theta) \le \exp\left(\frac{A}{\tilde{\Delta}}N^2\right)\right] \ge 1 - 4N\frac{\tilde{\Delta}}{\pi}, \tag{90}$$

where $\tilde{\Delta} > 0$ and A a positive numerical constant mildly depending on the group G. Using the bound $\Pr(A \cap B) \ge \Pr(A) + \Pr(B) - 1$ we obtain

$$\Pr_{g\sim\mu_G}\left[|\tilde{\mathcal{O}}[\Pi(g_{\theta_i})] - D_{g_0,g}(\theta_i)| \le \epsilon \exp\left(\frac{A}{\tilde{\Delta}}N^2\right)\right] > 1-\delta-8\Delta N^2 - 4N\frac{\tilde{\Delta}}{\pi}. \tag{91}$$

We finally use union bound lower to bound the probability that $\tilde{\mathcal{O}}$ is successful *for all* L equally spaced θ_i in $[1-\Delta, 1]$:

$$\Pr_{g\sim\mu_G}\left[\forall\theta_i |\tilde{\mathcal{O}}[\Pi(g_{\theta_i})] - D_{g_0,g}(\theta_i)| \le \epsilon \exp\left(\frac{A}{\tilde{\Delta}}N^2\right)\right] > 1 - L\left(\delta + 8\Delta N^2 + 4N\frac{\tilde{\Delta}}{\pi}\right). \tag{92}$$

If $L \approx \deg(D_{g_0,g}) = \Theta(N^2)$ the above evaluations of $\tilde{\mathcal{O}}$ can be used to recover polynomial $\tilde{P}_{g_0,g}$ passing through points $\{\theta_i, \tilde{\mathcal{O}}[\Pi(g_{\theta_i})]\}$ and having identical degree to $D_{g_0,g}$. By Eq. (91) and results of Coppersmith and Rivlin stated in Ref. [109] we know that [note that we set $L \approx \deg(D_{g_0,g}) = \Theta(N^2)$]

$$\begin{aligned}\max_{\theta\in[1-\Delta,1} \left|\tilde{P}_{g_0,g}(\theta) - D_{g_0,g}(\theta)\right| &\le \epsilon \exp\left(\frac{A}{\tilde{\Delta}}N^2\right)\exp[\Theta(N^2)] \\ &= \epsilon \exp\left(\frac{\Theta(N^2)}{\tilde{\Delta}}\right).\end{aligned} \tag{93}$$

Recall that by assumption and definition of Cayley path $D_{g_0,g}(0)$ encodes a (rescaled) #P-hard probability amplitude. Using Paturi lemma for the polynomial $\tilde{D}_{g_0,g}(\theta) - D_{g_0,g}(\theta)$ we finally obtain

$$\left|\tilde{D}_{g_0,g}(0) - D_{g_0,g}(0)\right| \le \epsilon \exp\left(\frac{\Theta(N^2)}{\tilde{\Delta}} + \Theta(N^2)(1+\Delta^{-1})\right). \tag{94}$$

To sum up, the initially assumed oracle $\mathcal{O}$ allows us to construct an efficient algorithm $\mathcal{A}$ that approximately computes #P-hard quantity $D_{g_0,g}(0) = Q_g(\theta)\,|\langle \mathbf{x}_0|\,\Pi(g_0)\,|\Psi_{\rm in}\rangle|^2$:

$$\Pr_{g\sim\mu_G}\{\left|\mathcal{A}[\Pi(g)] - D_{g_0,g}(0)|\right| \le \tilde{\epsilon}\} > 1 - BN^2\left(\delta + 8\Delta N^2 + 4N\frac{\tilde{\Delta}}{\pi}\right), \tag{95}$$

where $\tilde{\epsilon} = \epsilon\ \exp\left(\Theta(N^2)/\tilde{\Delta} + \Theta(N^2)(1+\Delta^{-1})\right)$, and $B > 0$ is a numerical constant. Success probability of the protocol to exceeds $\frac{1}{2}$ with the following: scaling

$$\Delta = \Theta(N^{-4}), \tilde{\Delta} = \Theta(N^{-3}). \tag{96}$$

From the result of Ref. [111] we have #P-hardness guarantees up to constant multiplicative error. Since for #P-hard quantity this such error implies additive error of magnitude at most $2^{-\Theta(N)}$. Therefore, by setting $\tilde{\epsilon} \le 2^{-\Theta(N)}$, which, by the virtue of Eq. (96), corresponds to scaling of the original error $\epsilon = \exp[-\Theta(N^6)]$ allowing to extrapolate to the hardness neighbored. ■

Remark 17: In the course of the proof of the above result we have realized an inadequate usage of the oracle in the reduction by Movassagh [15] [the author assumed that the oracle works as in Eq. (88) but without the necessary dependence on Δ]. Correction of the proof seems to give, in that case, worse than claimed tolerance for error $\epsilon = \exp[-\theta(N^{4.5})]$ (for the Google layout), which is still better then the one claimed here.

IX. EFFICIENT TOMOGRAPHY OF FERMIONIC LINEAR OPTICS

The tomography and certification of gates, i.e., the task of ensuring that the correct unitary was implemented, is vital for near-term quantum devices. However, it is often an inherently challenging problem due to exponential scaling of the number of parameters describing a general multiqubit quantum operation [112]. Here we show that the structure of FLO unitaries allows us to perform their tomography efficiently using resources scaling only polynomially with the system size. As passive fermionic gates form a subset of active FLO circuits, we focus only on the tomography of the latter ones, since from this also the tomography of passive circuit follows.

We use here again the Jordan-Wigner mapping between the d-qubit system and fermionic Fock space with d physical modes (see Sec. II), and define the following $2d$ pure states:

$$\begin{aligned} \left|+^p_X\right\rangle &= \mathbb{I}^{\otimes(p-1)} \otimes H \otimes \mathbb{I}^{\otimes(d-p)} |0\rangle^{\otimes d} \\ &= |0\rangle^{\otimes(p-1)} \otimes |+_X\rangle \otimes |0\rangle^{\otimes(d-p)}, \end{aligned} \tag{97}$$

$$\begin{aligned} \left|+^p_Y\right\rangle &= \mathbb{I}^{\otimes(p-1)} \otimes \tilde{H} \otimes \mathbb{I}^{\otimes(d-p)} |0\rangle^{\otimes d} \\ &= |0\rangle^{\otimes(p-1)} \otimes |+_Y\rangle \otimes |0\rangle^{\otimes(d-p)}, \end{aligned} \tag{98}$$

where $p = 1, \ldots, d$.

In terms of Majorana operators, one can write the density matrices of these states as

$$\rho_{2p-1} = \left|+^p_X\right\rangle\left\langle+^p_X\right| = \prod_{q=1}^{p-1}\left(\frac{\mathbb{I} + i m_{2q-1} m_{2q}}{2}\right)\left(\frac{\mathbb{I} + \prod_{q=1}^{p-1}(i m_{2q-1} m_{2q}) m_{2p-1}}{2}\right)\prod_{q=p+1}^{d}\left(\frac{\mathbb{I} + i m_{2q-1} m_{2q}}{2}\right), \tag{99}$$

$$\rho_{2p} = \left|+^p_Y\right\rangle\left\langle+^p_Y\right| = \prod_{q=1}^{p-1}\left(\frac{\mathbb{I} + i m_{2q-1} m_{2q}}{2}\right)\left(\frac{\mathbb{I} + \prod_{q=1}^{p-1}(i m_{2q-1} m_{2q}) m_{2p}}{2}\right)\prod_{q=p+1}^{d}\left(\frac{\mathbb{I} + i m_{2q-1} m_{2q}}{2}\right), \tag{100}$$

where, as before, $p = 1, \ldots, d$. Expanding these density matrices in terms of Majorana monomials [given in Eq. (D14)], we observe that for an arbitrary ρ_x ($x = 1, \ldots, 2d$) there is only one Majorana monomial of degree 1 appearing, namely m_x. Thus, considering the FLO evolved states $V\rho_x V^\dagger$, the degree 1 Majorana terms will be of the form [see Eq. (11)] $V m_x V^\dagger = \sum_{y=1}^{2d} O_{yx} m_y$, where $O \in \mathrm{SO}(2d)$ is the orthogonal matrix that encodes the FLO circuit V. In order to obtain arbitrary element of the orthogonal matrix O_{yx}, one needs only to insert the state ρ_x, evolve it with the FLO unitary V, and then measure the expectation value of m_y:

$$O_{yx} = \mathrm{tr}(m_y V \rho_y V^\dagger). \tag{101}$$

Measuring the expectation value of m_{2q-1} and m_{2q} amounts to measuring $Z_1 \cdots Z_{q-1} X_q$ and $Z_1 \cdots Z_{q-1} Y_q$, respectively. These can all be done, after a single layer of local base change operations, through usual computational basis measurements. The graphical presentation of our tomography scheme is given in Fig. 9. The following theorem show that the construction outlined above allows recovery of an unknown FLO circuit V efficiently in d, both in terms of the number of different setups needed for the implementation as well as in terms of sample complexity. Importantly, our results give rigorous recovery guarantees in the diamond norm, despite the presence of statistical fluctuations.

Theorem 8: (Efficient tomography of active FLO unitary channels). *Let V be an unknown active FLO circuit acting on d qubits. Consider the following estimation protocol*

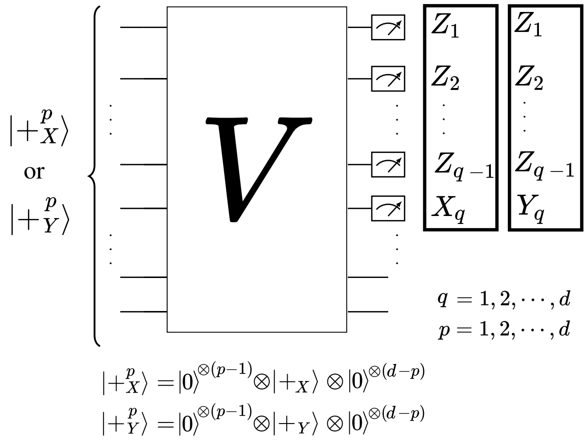

$|+^p_X\rangle$
or
$|+^p_Y\rangle$
V
Z_1
Z_2
Z_{q-1}
X_q
Z_1
Z_2
Z_{q-1}
Y_q
$q = 1, 2, \cdots, d$
$p = 1, 2, \cdots, d$
$|+^p_X\rangle = |0\rangle^{\otimes(p-1)} \otimes |+_X\rangle \otimes |0\rangle^{\otimes(d-p)}$
$|+^p_Y\rangle = |0\rangle^{\otimes(p-1)} \otimes |+_Y\rangle \otimes |0\rangle^{\otimes(d-p)}$

FIG. 9. Graphical presentation of the tomography protocol of an active FLO circuit V. A single step of the protocol consists of (i) preparation of $2d$ input states $\left|+^p_X\right\rangle$ and $\left|+^p_Y\right\rangle$ ($p = 1, \ldots, d$), (ii) transformation of the states via the circuit V, and (iii) for each of the $2d$ states measuring the operators $Z_1 Z_2 \cdots Z_{q-1} X_q$ and $Z_1 Z_2 \cdots Z_{q-1} Y_q$ ($q = 1, \ldots, d$). These operations are then repeated multiple times in order to gather sufficient statistics necessary to reconstruct the orthogonal matrix $O \in \mathrm{SO}(2d)$ that defines the unitary channel Φ_V associated to $V = \Pi_{\mathrm{act}}(O)$.

using the states ρ_x and observables m_y $(x,y \in [2d])$ and comprising of r independent experimental rounds. A single experimental round, say the k'th, consists of the following routines:

(a) For every pair $(x,y) \in [2d]^{\times 2}$: (i) prepare ρ_x as input state; (ii) evolve ρ_x via the circuit V; (iii) measure $V\rho_x V^\dagger$ using m_y obtaining outcome $m_{yx}^{(k)} \in \{-1,1\}$.

The outcomes of the k'th round are gathered in the $2d \times 2d$ matrix $M^{(k)}$ with entries $m_{yx}^{(k)}$. After r rounds, define $\hat{M}_r := 1/r\sum_{k=1}^r M^{(k)}$ as the sample average of matrices $M^{(k)}$. Then, let $\hat{O}_r \in SO(2d)$ be defined as the orthogonal matrix appearing in the polar decomposition of $\hat{M}_r$ (i.e., $\hat{M}_r = O_r P$, where P is a semidefinite real $2d \times 2d$ matrix). Finally, set $\hat{V} := \Pi_{\mathrm{act}}(\hat{O}_r)$ as the estimator of the circuit V after r rounds of the protocol.

Assume that all routines in the protocol are implemented perfectly. Furthermore, let $\delta \in (0,1)$ be fixed and let Φ_V and $\Phi_{\hat{V}}$ be the unitary channels defined by the active FLO circuits V and $\hat{V}$, respectively. Then, for the number of rounds satisfying

$$r \geq \frac{28d^3}{\epsilon^2}\log\left(\frac{4d}{\delta}\right), \tag{102}$$

the protocol outputs a FLO circuit $\hat{V}$ such that $\|\Phi_V - \Phi_{\hat{V}}\|_\diamond \leq \epsilon$ with probability at least $1-\delta$.

Remark 18: We believe that it possible to improve the sampling complexity and the number of quantum circuits needed for the tomography of an unknown FLO unitary V. Moreover, we expect that our proof technique can also be used for the quantum process tomography of general fermionic Gaussian channels.

There are three key difficulties that need to be circumvented in order to establish the above result. The first one is related to the fact that, by the virtue of Eq. (101), the protocol estimates an orthogonal matrix $O \in \mathrm{SO}(2d)$ not the circuit V or the associated d-qubit channel Φ_V. The following lemma, proved in Appendix F, allows us to connect operator-norm distance between elements of the orthogonal group with the diamond norm between the corresponding quantum channels (this result can be viewed as a fermionic version of the analogous stability result proved by Arkhipov for standard boson sampling [78]).

Lemma 6: (Stability of the active FLO representation). *Consider two elements of the orthogonal group, $O, O' \in SO(2d)$, and let V and V′ be the corresponding active FLO unitaries, i.e., $V = \Pi_{\mathrm{act}}(O)$ and $V' = \Pi_{\mathrm{act}}(O')$. Furthermore, let Φ_V and $\Phi_{V'}$ be the unitary channels defined by V and V′, respectively. Then the following inequality is satisfied:*

$$\|\Phi_V - \Phi_{V'}\|_\diamond \leq 2d\|O - O'\|. \tag{103}$$

The second technical issue arises because the sample-average matrices $\hat{M}_s$ appearing in the protocol are not necessarily orthogonal. For this reason we use the (real) polar decomposition in order to get an orthogonal matrix from $\hat{M}_s$. The lemma below gives an upper bound for the possible operator-norm error that can result from this procedure.

Lemma 7: (Operator-norm stability of the real polar decomposition [113]). *Let O be orthogonal matrix $n \times n$. Let Δ be $n \times n$ real matrix such that $\|\Delta\| \leq 1$. Let O_Δ be the orthogonal transformation appearing in the polar decomposition of $O + \Delta A$ (i.e., $O + \Delta = O_{O+\Delta}P$ for a semidefinite real matrix P). We then have the following inequality:*

$$\|O - O_\Delta\| \leq \|\Delta\|. \tag{104}$$

The above lemma follows as a direct corollary of Theorem 2.3 in Ref. [113].

The last technical ingredient needed for the proof of Theorem 8 is the following matrix concentration bound, which allows control of the magnitude of statistical fluctuations incurred in our scheme.

Lemma 8: (Matrix Bernstein inequality [114]). *Let $S^{(1)}, \ldots, S^{(r)}$ be independent, centered real $n \times n$ random matrices with uniformly bounded operator norm, i.e., for all $k \in [r]$*

$$\mathbb{E}S^{(k)} = 0,\ \|S^{(k)}\| \leq L. \tag{105}$$

Assume furthermore that the entries of each $S^{(k)}$ are independently distributed with a variance upper bounded by a constant, $\mathrm{Var}(S_{ij}^{(k)}) \leq c$.

We then have the following concentration inequality valid for arbitrary $\tau > 0$:

$$\Pr\left(\left\|\frac{1}{r}\sum_{k=1}^r S^{(k)}\right\| \geq \tau\right) \leq 2n\exp\left(-\frac{r\tau^2}{2(nc + \frac{L}{3}\tau)}\right). \tag{106}$$

A more general version of the above inequality (that does not require independently distributed entries of matrices $S^{(k)}$) can be found in Theorem 1.6.2 from Ref. [114].

Proof. Let us remark first that our tomography protocol was defined such that the matrices $M^{(k)}$ originating from different rounds k are independent from each other, and for fixed k also their entries $m_{yx}^{(k)}$ are independent. Furthermore, by virtue of Eq. (101), we have

$$\mathbb{E}m_{yx}^{(k)} = O_{yx}, \tag{107}$$

where $O \in \mathrm{SO}(2d)$ is an orthogonal matrix corresponding to the circuit V. We now apply Lemma 8 to the sequence of $2d \times 2d$ matrices $\Delta^{(k)} := M^{(k)} - O$. From definition matrix elements of $\Delta^{(k)}$ satisfy $|\Delta_{yx}^{(k)}| \leq 2$. From this and the fact that $m_{yx}^{(k)} \in \{-1, 1\}$ it easily follows that

$$\|\Delta^{(k)}\| \leq 4d, \ \mathrm{Var}(\Delta_{yx}^{(k)}) \leq 1. \tag{108}$$

Inserting these estimates in Eq. (106) (and noting that $n = 2d$) gives

$$\Pr\left(\left\|\frac{1}{r}\sum_{k=1}^{r}\Delta^{(k)}\right\| \geq \tau\right) \leq 4d \exp\left(-\frac{r\tau^2}{4d(1+\frac{2}{3}\tau)}\right). \tag{109}$$

Recalling that $\hat{M}_r = 1/r\sum_{k=1}^{r} M^{(k)}$, using the definition of $\Delta^{(k)}$, and assuming that $\tau < 1$ (in what follows we see that we can introduce this constraint without the loss of generality) we obtain

$$\Pr\left(\left\|\hat{M}_r - O\right\| \leq \tau\right) \geq 1 - 4d \exp\left(-\frac{r\tau^2}{7d}\right). \tag{110}$$

We therefore know that, provided r is high enough, the sample average $\hat{M}^{(k)}$ approximates matrix O in operator norm. Applying Lemma 7 to $\hat{O}_r$, i.e., to the orthogonal part of the polar decomposition of $\hat{M}_r$ [this corresponds to setting $\Delta = \hat{M}_r - O$ in Eq. (104)], we obtain

$$\Pr\left(\left\|\hat{O}_r - O\right\| \leq \tau\right) \geq 1 - 4d \exp\left(-\frac{r\tau^2}{7d}\right). \tag{111}$$

Recalling that $\hat{V} = \Pi_{\mathrm{act}}(\hat{O})$ and $V = \Pi_{\mathrm{act}}(O)$ and invoking Lemma 6 we finally arrive to

$$\Pr\left(\left\|\Phi_{\hat{V}} - \Phi_V\right\|_{\diamond} \leq 2d\,\tau\right) \geq 1 - 4d \exp\left(-\frac{r\tau^2}{7d}\right). \tag{112}$$

We conclude the proof by setting $\epsilon := 2d\,\tau$ and noting that Eq. (102) follows from requiring that the right-hand side of Eq. (112) is larger than $1 - \delta$. ■

ACKNOWLEDGMENTS

We would like to thank Ramis Movassagh for explanations regarding his work. M.O. and N.D. acknowledge support by the Foundation for Polish Science through the TEAM-NET project (contract no. POIR.04.04.00-00-17C1/18-00). Z.Z. was supported by the NKFIH through the Quantum Technology National Excellence Program (Project No. 2017-1.2.1-NKP-2017-00001), the Grants No. K124152, No. FK135220, No. KH129601, and No. K120569 and the Quantum Information National Laboratory of Hungary. M.M. was supported by the ARC Centre of Excellence for Quantum Computation and Communication Technology (CQC2T), (Project No. CE170100012) and a scholarship top-up and extension from the Sydney Quantum Academy.

APPENDIX

We collect here technical results that are used in the main part of the paper. Some of the results stated here can be of independent interest for further works on quantum information processing with fermions.

APPENDIX A: DECOMPOSITION OF PASSIVE AND ACTIVE FLO UNITARIES INTO TWO-QUBIT GATES

Here we provide the derivation of the decomposition of arbitrary passive and active FLO gates into two-qubit gates with layouts depicted in Fig. 4, which was also studied in Refs. [38–40]. For passive bosonic linear optics the analogous decompositions were discussed in Refs. [5,6]. The way to obtain these results is to consider the standard decomposition of $\mathrm{U}(d)$ and $\mathrm{SO}(2d)$ elements into so-called Givens rotations and then apply the appropriate FLO representations Π_{pas} and Π_{act} on this decomposition, as we explain below. For simplicity, we assume that d is even, which is also the relevant case for our paper.

The (nearest-neighbor) Givens rotations $G^k(\alpha, \varphi) \in \mathrm{U}(d)$ $(k = 1, \ldots, d{-}1)$ have the form

$$G^{(k)}(\alpha,\varphi) = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & e^{i\varphi}\cos(\alpha) & -\sin(\alpha) & \cdots & 0 \\ 0 & \cdots & e^{i\varphi}\sin(\alpha) & \cos(\alpha) & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix}, \tag{A1}$$

where only the 2×2 block consisting of the entries with row and column indices k and $k+1$ are nontrivial. A general element $U \in \mathrm{U}(d)$ can then be decomposed into Givens rotations in different ways, we consider two of these (also discussed in Ref. [5]). In the first decomposition one applies alternatingly (d number of times) a series of Givens rotations $G^{(k)}$ with odd and even k indices, and finally a diagonal unitary $T = \mathrm{diag}(e^{i\kappa_1}, e^{i\kappa_2}, \dots, e^{i\kappa_d})$, i.e.,

$$U = TB_{d/2}A_{d/2}\dots B_2A_2B_1A_1, \tag{A2}$$

$$A_j = \prod_{\substack{k\in[d-1]\\ k \text{ odd}}} G^{(k)}(\alpha_{(k,j)}, \varphi_{(k,j)}),$$
$$B_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}} G^{(k)}(\beta_{(k,j)}, \nu_{(k,j)}),\ j \in [d/2]. \tag{A3}$$

Also in the second decomposition one applies alternatingly a series of Givens rotations $G^{(k)}$ with odd and even k indices, however, this time there is $2d-1$ such layers and in the ℓth layer there are only Givens rotations up to index $(d-|\ell-d|)$, and finally there is again a diagonal unitary $T' = \mathrm{diag}(e^{i\kappa'_1}, e^{i\kappa'_2}, \dots, e^{i\kappa'_d})$, i.e.,

$$U = T'A'_dB'_{d-1}\dots B'_2A'_2B'_1A'_1, \tag{A4}$$

$$A'_j = \prod_{\substack{k=1\\ k \text{ odd}}}^{d-|2j-d|} G^{(k)}(\gamma_{(k,j)}, \tau_{(k,j)}),$$
$$B'_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}}^{d-|2j-d|} G^{(k)}(\delta_{(k,j)}, \sigma_{(k,j)}),\ j \in [d/2]. \tag{A5}$$

Note that both decompositions use the same number of elementary Givens rotations.

Now, given an arbitrary passive FLO transformation $V = \Pi_{\mathrm{pas}}(U)$, with $U \in \mathrm{U}(d)$, we can use the fact that Π_{pas} is a representation (and thus a homomorphism) and apply it to the decompositions of Eqs. (A2) and (A4). We obtain

$$V = \Pi_{\mathrm{pas}}(U) = \Pi_{\mathrm{pas}}(T)L_{d/2}K_{d/2}\dots L_2K_2L_1K_1, \tag{A6}$$

$$K_j = \prod_{\substack{k\in[d-1]\\ k \text{ odd}}} \Pi_{\mathrm{pas}}[G^{(k)}(\alpha_{(k,j)}), \varphi_{(k,j)}],$$
$$L_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}} \Pi_{\mathrm{pas}}[G^{(k)}(\beta_{(k,j)}, \nu_{(k,j)})],\ j \in [d/2], \tag{A7}$$

and similarly

$$V = \Pi_{\mathrm{pas}}(U) = \Pi_{\mathrm{pas}}(T')K'_dL'_{d-1}\dots L'_2K'_2L'_1K'_1, \tag{A8}$$

$$K'_j = \prod_{\substack{k=1\\ k \text{ odd}}}^{d-|2j-d|} \Pi_{\mathrm{pas}}[G^{(k)}(\gamma_{(k,j)}, \tau_{(k,j)})],$$
$$L'_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}}^{d-|2j-d|} \Pi_{\mathrm{pas}}[G^{(k)}(\delta_{(k,j)}, \sigma_{(k,j)})],\ j \in [d/2]. \tag{A9}$$

Using the definition of Π_{pas} and the Jordan-Wigner correspondence between fermions and qubits systems, we have that

$$\Pi_{\mathrm{pas}}[\mathrm{diag}(e^{i\alpha}, e^{i\alpha_2}, \dots, e^{i\alpha_d})] = e^{i\alpha_1 Z} \otimes e^{i\alpha_1 Z} \otimes \dots \otimes e^{i\alpha_d Z}, \tag{A10}$$

$$\Pi_{\mathrm{pas}}[G^{(k)}(\alpha_1, \alpha_2)] = \mathbb{I}^{\otimes k-1} \otimes (e^{-i\alpha_1 Z/2} \otimes e^{i\alpha_1 Z/2}) \times e^{i\alpha_2 (X\otimes X + Y\otimes Y)/2} \otimes \mathbb{I}^{\otimes d-k-1}. \tag{A11}$$

Thus, Eqs. (A16) and (A18) provide exactly the brickwall and triangle decomposition of Fig. 4.

Let us now turn to the decomposition of an arbitrary active FLO gate $V = \Pi_{\mathrm{act}}(O)$ $[O \in \mathrm{SO}(2d)]$. An orthogonal matrix O can be decomposed into a sequence of real Givens rotations $G^{(k)}(\alpha) := G^{(k)}(\alpha, 0) \in \mathrm{SO}(2d)$ analogously to the decompositions of a unitary [Eqs. (A2) and (A4)]. One can apply alternatingly (d number of times) a series of real Givens rotations $G^{(k)}$ with odd and even k indices, and finally a diagonal orthogonal matrix $S = \mathrm{diag}(s_1, s_2, \dots, s_d)$ (with $s_i \in \{1, -1\}$ and $\prod_{i=1}^{2d} = 1$), i.e.,

$$U = SD_{d/2}C_{d/2}\dots D_2C_2D_1C_1, \tag{A12}$$

$$C_j = \prod_{\substack{k\in[d-1]\\ k \text{ odd}}} G^{(k)}(\alpha_{(k,j)}),$$
$$D_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}} G^{(k)}(\beta_{(k,j)}),\ j \in [d/2]. \tag{A13}$$

Alternatively, one can apply alternatingly a series of Givens rotations $G^{(k)}$ with odd and even k indices with $2d-1$ layers and in the ℓth layer there are only real Givens rotations up to index $(d-|\ell-d|)$, and finally there is again a diagonal matrix with signs $S' = \mathrm{diag}(s'_1, s'_2, \dots, s'_d)$, i.e.,

$$U = S'C'_dD'_{d-1}\dots D'_2C'_2D'_1C'_1, \tag{A14}$$

$$C'_j = \prod_{\substack{k=1\\ k \text{ odd}}}^{d-|2j-d|} G^{(k)}(\gamma_{(k,j)}),$$
$$D'_j = \prod_{\substack{k\in[d-1]\\ k \text{ even}}}^{d-|2j-d|} G^{(k)}(\delta_{(k,j)}),\ j \in [d/2]. \tag{A15}$$

Given an arbitrary active FLO transformation $V = \Pi_{\mathrm{act}}(O)$, with $O \in \mathrm{SO}(2d)$, we can use the fact that Π_{act} is a projective representation (and thus a projective homomorphism) and apply it to the decompositions of Eqs. (A12) and (A14), obtaining upto irrelevant signs

$\sigma, \sigma' \in \{1, -1\}$ that

$$V = \Pi_{\text{act}}(O) = \sigma\, \Pi_{\text{act}}(S) F_{d/2} E_{d/2} \ldots F_2 E_2 F_1 E_1, \quad \text{(A16)}$$

$$E_j = \prod_{\substack{k \in [d-1] \\ k \text{ odd}}} \Pi_{\text{act}}[G^{(k)}(\alpha_{(k,j)})],$$

$$F_j = \prod_{\substack{k \in [d-1] \\ k \text{ even}}} \Pi_{\text{act}}[G^{(k)}(\beta_{(k,j)})], \; j \in [d/2], \quad \text{(A17)}$$

and

$$V = \Pi_{\text{act}}(O) = \sigma' \Pi_{\text{pas}}(S') F'_d E'_{d-1} \ldots F'_2 E'_2 F'_1 E'_1, \quad \text{(A18)}$$

$$E'_j = \prod_{\substack{k=1 \\ k \text{ odd}}}^{d-|2j-d|} \Pi_{\text{act}}[G^{(k)}(\gamma_{(k,j)})],$$

$$L'_j = \prod_{\substack{k \in [d-1] \\ k \text{ even}}}^{d-|2j-d|} \Pi_{\text{act}}[G^{(k)}(\delta_{(k,j)})], \; j \in [d/2]. \quad \text{(A19)}$$

Using the definition of Π_{pas} and the Jordan-Wigner correspondence between fermions and qubits systems, we have that

$$\Pi_{\text{act}}(S) = \pm X^{s_1} Y^{s_2} \otimes X^{s_3} Y^{s_4} \otimes \ldots \otimes X^{s_{2d-1}} Y^{s_{2d}}, \quad \text{(A20)}$$

and

$$\Pi_{\text{pas}}[G^{(k)}(\alpha)] = e^{-\alpha\, m_k m_{k+1}} = \begin{cases} \mathbb{I}^{\otimes(\ell-1)} \otimes e^{i\alpha Z_\ell} \otimes \mathbb{I}^{\otimes(d-\ell)} & \text{if } k = 2\ell \text{ is even} \\ \mathbb{I}^{\otimes(\ell-1} \otimes e^{i\alpha X_\ell X_{\ell+1}} \otimes \mathbb{I}^{\otimes(d-\ell-1)} & \text{if } k = 2\ell+1 \text{ is odd.} \end{cases} \quad \text{(A21)}$$

Thus, the circuits would resemble the brickwall and layouts, however with depths $2d$ and $(4d-1)$ on $2d$ Majorana lines and not of depth d and $2d-1$ on d qubit lines, see Fig. 10. (In circuits with Majorana lines, the lines represent individual operators and gates between two Majorana lines are unitaries that is composed only of the corresponding two Majorana operators [55].) However, we can make some simplifications by merging gates as shown in Fig. 10: in the middle of the circuit we can merge four two-qubit gates (corresponding to four Givens rotations) of the form $e^{i\alpha_1 X\otimes X}$ $(e^{i\alpha_2 Z} \otimes e^{i\alpha_3 Z})$ $e^{i\alpha_4 X \otimes X}$ and these are equal to gates of the form $D_{\text{act}}(\{\beta_i\}) = (e^{i\beta_5 Z/2} \otimes e^{i\beta_6 Z/2})\, e^{i(\beta_3 X\otimes X + \beta_4 Y \otimes Y)/2}\, (e^{i\beta_1 Z/2} \otimes e^{i\beta_2 Z/2})$, where the β_i's have to be chosen to satisfy

$$\cos(\alpha_1 + \alpha_4)\cos(\alpha_2 - \alpha_3) = \cos(\theta_2)\cos(\theta_1 + \theta_3),$$
$$\times \sin(\alpha_1 + \alpha_4)\cos(\alpha_2 - \alpha_3) = \sin(\theta_2)\cos(\theta_1 - \theta_3), \quad \text{(A22)}$$

$$\cos(\alpha_1 - \alpha_4)\sin(\alpha_2 - \alpha_3) = \cos(\theta_2)\sin(\theta_1 + \theta_3),$$
$$\times \cos(\alpha_1 - \alpha_4)\sin(\alpha_2 + \alpha_3) = \cos(\theta_5)\sin(\theta_4 + \theta_6), \quad \text{(A23)}$$

$$\cos(\alpha_1 + \alpha_4)\cos(\alpha_2 + \alpha_3) = \cos(\theta_5)\cos(\theta_4 + \theta_6),$$
$$\times \sin(\alpha_1 + \alpha_4)\cos(\alpha_2 + \alpha_3) = \sin(\theta_5)\cos(\theta_4 - \theta_6), \quad \text{(A24)}$$

where we use the notations $\theta_1 = \beta_1 - \beta_2$, $\theta_2 = \beta_3 - \beta_4$, $\theta_3 = \beta_5 - \beta_6$, $\theta_4 = \beta_1 + \beta_2$, $\theta_5 = \beta_3 + \beta_4$, $\theta_6 = \beta_5 + \beta_6$. At the edges of the circuit we may just either have to join additional local Z rotations to the merged gates, thus it can be again expressed as $D_{\text{act}}(\{\beta_i\})$, or it is already of the form of $D_{\text{act}}(\{\beta_i\})$. In this way, we obtain exactly the brickwall and triangle decomposition of Fig. 4 with two-qubit gates of the form of $D_{\text{act}}(\{\beta_i\})$.

APPENDIX B: PROOF OF THEOREM 3

Proof. We consider in parallel active and passive FLO circuits. For passive FLO we have $\mathcal{H} = \mathcal{H}_{\text{pas}}$ and $\nu = \nu_{\text{pas}}$

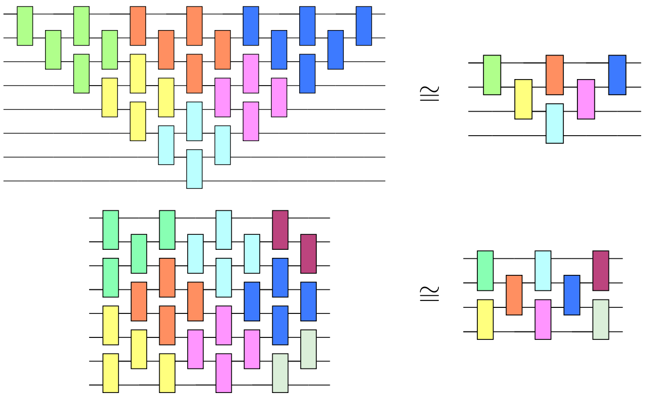


FIG. 10. Decomposition of an arbitrary $V = \Pi_{\text{act}}(O)$ using Majorana-line (left) and qubit-line (right) circuit pictures. The represented Givens rotations can be merged (identical colors depicting the merged rotations) giving rise to a layout of Fig. 4, with two-qubit gates of type $D_{\text{act}}(\{\beta_i\})$.

while for active FLO we have $\mathcal{H}_{\mathrm{act}}$ and $\nu = \nu_{\mathrm{act}}$. With the fixed input $|\Psi_{\mathrm{in}}\rangle = |\Psi_4\rangle^{\otimes N}$, we write $p_{\mathbf{x}}(V) = |\langle \mathbf{x}|V|\Psi_{\mathrm{in}}\rangle|^2$ for the probability of outcome $\mathbf{x}$ [we assume that $|x\rangle \in \mathcal{H}$], and $p(V)$ for the output probability distribution of a circuit V. Suppose that there exists a classical sampler $\mathcal{C}$ that performs fermion sampling for a fixed but arbitrary FLO circuit V, and denote by $q(V)$ the distribution from which $\mathcal{C}$ samples. Then for a given $\mathbf{x}$, by Stockmeyer's approximate counting algorithm [53], a BPP^{NP} machine with an oracle access to $\mathcal{C}$ can produce a multiplicative estimates $\tilde{q}_{\mathbf{x}}(V)$ of $q_{\mathbf{x}}(V)$ such that

$$|q_{\mathbf{x}}(V) - \tilde{q}_{\mathbf{x}}(V)| \leq \frac{q_{\mathbf{x}}}{\mathrm{poly}(N)} \tag{B1}$$

for every $\mathbf{x}$. We show that $\tilde{q}_{\mathbf{x}}(V)$ is also close to $p_{\mathbf{x}}(V)$ for most $\mathbf{x}$ and V that anticoncentrate. Judiciously applying the triangle inequality, we have that

$$\begin{aligned}
&|p_{\mathbf{x}}(V) - \tilde{q}_{\mathbf{x}}(V)| \\
&\quad\leq |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| + |q_{\mathbf{x}}(V) - \tilde{q}_{\mathbf{x}}(V)| && \text{(B2)}\\
&\quad\leq |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| + \frac{q_{\mathbf{x}}(V)}{\mathrm{poly}(N)} && \text{(B3)}\\
&\quad\leq |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| + \frac{|p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| + p_{\mathbf{x}}(V)}{\mathrm{poly}(N)} && \text{(B4)}\\
&\quad= \frac{p_{\mathbf{x}}(V)}{\mathrm{poly}(N)} + |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)|\left(1 + \frac{1}{\mathrm{poly}(N)}\right). && \text{(B5)}
\end{aligned}$$

Given that the distributions $p(V)$ and $q(V)$ are ϵ close in the l_1 norm, particular probabilities $p_{\mathbf{x}}(V)$ and $q_{\mathbf{x}}(V)$ must be exponentially close for most $\mathbf{x}$. This statement is made precise using Markov's inequality: for a non-negative random variable X and $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}X}{a}. \tag{B6}$$

Setting $X = |p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)|$ and $a = \epsilon/(|\mathcal{H}|\delta)$, (the probability is over the outcomes $\mathbf{x}$ which is distributed uniformly over $\mathcal{H}$, see Remark 8)

$$\begin{aligned}
&\Pr_{\mathbf{x}\sim\mathrm{unif}(\mathcal{H})}\left(|p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)| \geq \frac{\epsilon}{|\mathcal{H}|\delta}\right)\\
&\quad\leq \frac{\mathbb{E}_{\mathbf{x}\sim\mathrm{unif}(\mathcal{H})}[|p_{\mathbf{x}}(V) - q_{\mathbf{x}}(V)|]|\mathcal{H}|\delta}{\epsilon} \leq \delta. \qquad \text{(B7)}
\end{aligned}$$

Combining the probability bound with the inequality (B5), we have that with probability at least $1-\delta$ over random $\mathbf{x} \sim \mathrm{unif}(\mathcal{H})$ we have

$$|p_{\mathbf{x}}(V) - \tilde{q}_{\mathbf{x}}(V)| < \frac{p_{\mathbf{x}}(V)}{\mathrm{poly}(N)} + \frac{\epsilon}{|\mathcal{H}|\delta}\left(1 + \frac{1}{\mathrm{poly}(N)}\right). \tag{B8}$$

To turn the above additive upper bound to a multiplicative one, we use the anticoncentration property (Theorem 1), which lets us replace $1/|\mathcal{H}|$ by an upper bound $p_{\mathbf{x}}(V)/\alpha$ with probability $(1-\alpha)^2/C$.

In order to do so, we must consider the joint probability of $(V, \mathbf{x})$ as described in Remark 8. Let A be the event that $p_{\mathbf{x}}(V)$ and $q_{\mathbf{x}}(V)$ for a fixed V are exponential close due to Markov's inequality, and B be the event that the distribution $p(V)$ anticoncentrates. The probability of both "good events" happening is lower bounded by $\Pr(A \cap B) \geq \max\{0, \Pr(A) + \Pr(B) - 1\}$. That is, if we denote by $\mathcal{A}(V, \mathbf{x})$ an event that

$$\begin{aligned}
&|p_{\mathbf{x}}(V) - \tilde{q}_{\mathbf{x}}(V)|\\
&\quad< p_{\mathbf{x}}(V)\left[\frac{1}{\mathrm{poly}(N)} + \frac{\epsilon}{\alpha\delta}\left(1 + \frac{1}{\mathrm{poly}(N)}\right)\right], \qquad \text{(B9)}
\end{aligned}$$

we have that

$$\Pr_{V\sim\nu, \mathbf{x}\sim\mathrm{unif}(\mathcal{H})}[\mathcal{A}(V, \mathbf{x})] > \frac{(1-\alpha)^2}{C} - \delta, \tag{B10}$$

which can be simplified by using the hiding property described in Lemma 1. The property implies that $p_x(V) = p_{\mathbf{x}_0}(V_{\mathbf{x}})$ and $\tilde{q}_{\mathbf{x}}(V) = \tilde{q}_{\mathbf{x}_0}(V_{\mathbf{x}})$ so that

$$\Pr_{V\sim\nu, \mathbf{x}\sim\mathrm{unif}(\mathcal{H})}[\mathcal{A}(V, \mathbf{x})] = \mathop{\mathbb{E}}_{\mathbf{x}\sim\mathrm{unif}(\mathcal{H})}\left(\Pr_{V\sim\nu}[\mathcal{A}(V_{\mathbf{x}}, \mathbf{x}_0)]\right). \tag{B11}$$

Moreover, from the invariance of the Haar measure it follows that for every $|\mathbf{x}\rangle \in \mathcal{H}$, $V_{\mathbf{x}}$ is distributed in the same way as V. Consequently,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{x}\sim\mathrm{unif}(\mathcal{H})}\left(\Pr_{V\sim\nu}[\mathcal{A}(V_{\mathbf{x}}, \mathbf{x}_0)]\right) &= \mathop{\mathbb{E}}_{\mathbf{x}\sim\mathrm{unif}(\mathcal{H})}\left(\Pr_{V\sim\nu}[\mathcal{A}(V, \mathbf{x}_0)]\right)\\
&= \Pr_{V\sim\nu}[\mathcal{A}(V, \mathbf{x}_0)]. \qquad \text{(B12)}
\end{aligned}$$

We finally obtain that for every $\mathbf{x}_0$,

$$\Pr_{V\sim\nu}\left\{|p_{\mathbf{x}_o}(V) - \tilde{q}_{\mathbf{x}_o}(V)| < p_{\mathbf{x}_o}(V)\left[\frac{1}{\mathrm{poly}(N)} + \frac{\epsilon}{\alpha\delta}\left(1 + \frac{1}{\mathrm{poly}(N)}\right)\right]\right\} > \frac{(1-\alpha)^2}{C} - \delta. \tag{B13}$$

Following Ref. [11] and requiring a constant ϵ and relative error $\epsilon/(\alpha\delta)$ (see Remark 19) we may set, for instance,

$$\alpha = \frac{1}{2}, \qquad \delta = \frac{(1-\alpha)^2}{2C} = \frac{1}{8C}, \qquad \epsilon = \frac{\alpha\delta}{4} = \frac{1}{64C}, \tag{B14}$$

Stockmeyer's algorithm is able to output $[1/4 + o(1), 1/(8C)]$-multiplicative approximates of the output probabilities for $1/(8C)$ fraction of the (passive or active, with constant C_{pas} or C_{act}, respectively) FLO circuits V if there is a classical machine that approximately sample from $p_{\mathbf{x}}(V)$ for any FLO circuit V within the l_1 distance $1/(64C)$. ■

Remark 19: Of the three parameters ϵ, δ, and α, the l_1 distance ϵ and the relative error $\epsilon/(\alpha\delta)$ are typically assumed to be constant $[1/4 + o(1)$ for the latter$]$ in quantum advantage proposals [12,18,19,52]. In which case δ is also a constant, and then one optimizes for the constant α. However, one may allow ϵ to decay inverse polynomially in the size of the system while retaining a sensible notion of simulation by the sampling task [10,20]. Doing so allows a more plausible (weaker) average-case hardness assumption but the sampling task becomes more demanding.

APPENDIX C: TV DISTANCE BETWEEN HAAR MEASURE AND ITS CAYLEY-PATH DEFORMATIONS

In this part we prove Lemma 2 that upper bounds total variation distance between the Haar measures μ_G and their deformed versions μ_G^θ, for $G = \mathrm{U}(d)$ and $G = \mathrm{SO}(2d)$. In what follows we use the notions and notation established in Sec. VII.

Lemma 9: *Let G be equal to $U(d)$ or $SO(2d)$. Let $g_0 \in G$ be a fixed element in G. Let $g \sim \mu_G$ an let $g_\theta = g_0 F_\theta(g)$, for $\theta \in [0,1]$ and $F_\theta : G \rightarrow G$ defined in Eq. (47). Let now μ_G^θ denote the induced measure according to which g_θ? is distributed. Assume furthermore that $\theta \leq 1 - \Delta$, for $\Delta >$ 0. We then have*

$$\begin{aligned} \left\| \mu_{U(d)} - \mu_{U(d)}^\theta \right\|_{\text{TVD}} &\leq d^2\Delta/2, \\ \left\| \mu_{SO(2d)} - \mu_{SO(2d)}^\theta \right\|_{\text{TVD}} &\leq d^2\Delta/2. \end{aligned} \tag{C1}$$

Proof. By the block diagonalization previously discussed, the TVD can be computed in terms of an integral on maximal torus $\mathbb{T}$ of G:

$$\left\| \mu_G - \mu_G^\theta \right\|_{\text{TVD}} = \frac{1}{2} \int_{\mathbb{T}} d\boldsymbol{\varphi} \left| \mu_G(\boldsymbol{\varphi}) - \mu_G^\theta(\boldsymbol{\varphi}) \right|. \tag{C2}$$

μ_G is the distribution for generic $g \in G$ of the "generalized eigenvalues" $\boldsymbol{\phi} := (\phi_1, \dots, \phi_d)$, quantities that are invariant under conjugation by any element of G, and it is given by the celebrated Weyl's integration formulas: ■

Fact 1: [Weyl's integration formula for $\mathrm{U}(d)$ and $\mathrm{SO}(d)$].

$$\begin{aligned} \mu_{\mathrm{U}(d)}(\boldsymbol{\phi}) &= \frac{1}{d!(2\pi)^d} \prod_{1 \leq j,k \leq d} \left| e^{i\phi_k} - e^{i\phi_j} \right|^2, \\ \mu_{\mathrm{SO}(2d)}(\boldsymbol{\phi}) &= \frac{2}{d!(2\pi)^d} \prod_{1 \leq j,k \leq d} 4[\cos(\phi_k) - \cos(\phi_j)]^2. \end{aligned} \tag{C3}$$

To upper bound the TVD, we use the fact that the Haar measure $\mu_G(\boldsymbol{\phi})$ is induced from $\mu_G^\theta(\boldsymbol{\varphi})$ by the inverse map F_θ^{-1} to compute $\mu_G(\boldsymbol{\varphi})$ by Fact 2 below. In particular, we show that the two measures μ_G and μ_G^θ expressed in the same coordinates $\boldsymbol{\varphi}$ are proportional to each other and bound the proportionality constant.

Fact 2: (Transport of measure). Let M and N be d-dimensional smooth manifolds with local coordinates $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_d)$ and $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots)$, μ a measure on M, and $F : M \rightarrow N$ a smooth map. Then

$$\tilde{\mu} = \mu \circ F^{-1} \tag{C4}$$

is a measure on N transported under F, where F^{-1} denotes the preimage of F. In particular, for any measurable set $A \subset N$,

$$\tilde{\mu}(A) = \mu[F^{-1}(A)] := \int_{F^{-1}(A)} \mu(\boldsymbol{\phi}) d\boldsymbol{\phi}. \tag{C5}$$

Explicitly, since the manifolds are locally Euclidean, $\tilde{\mu}(A)$ has an expression in terms of the Jacobian:

$$\tilde{\mu}(A) = \int_A \mu[F^{-1}(d\boldsymbol{\varphi})] \left| DF^{-1}(\boldsymbol{\varphi}) \right|, \tag{C6}$$

where $\left| DF^{-1}(\boldsymbol{\varphi}) \right|$ is the Jacobian, which by the inverse function theorem

$$\left| DF^{-1}(\boldsymbol{\varphi}) \right| = \left| DF[F^{-1}(\boldsymbol{\varphi})] \right|^{-1}. \tag{C7}$$

Remark 20: Since $F^{-1}(\boldsymbol{\varphi}) = \boldsymbol{\phi}$, the formula can be interpreted as a change of variable from $\boldsymbol{\varphi}$ to $\boldsymbol{\phi}$. In our case, $F_\theta^{-1} : \mathbb{T}(\boldsymbol{\phi}) \rightarrow \mathbb{T}(\boldsymbol{\varphi})$ plays the role of F. Equation (C10) is precisely the change of variable induced by F_θ^{-1}.

As an intermediate step, let us derive the explicitly change-of-variable formula from $\boldsymbol{\varphi}$ to $\boldsymbol{\phi}$. By applying to

$$\exp(\varphi_j \tilde{X}_j) = F_\theta\, [\exp(\phi_j \tilde{X}_j)], \tag{C8}$$

[recalling that $\tilde{X}$ are generators of the maximal torus for U(d) and SO($2d$)] the identity

$$f^{-1}[\exp(\phi_j \tilde{X}_j)] = -\tan(\phi_j/2)\tilde{X}_j, \tag{C9}$$

which can be verified by explicitly computing the Cayley transform (44) of at most a 2×2 matrix, we obtain the change-of-variable formula:

$$\varphi = 2\tan^{-1}[\theta\tan(\phi/2)]. \tag{C10}$$

Now we compute and bound (C6)

$$\mu_G(A) = \int_A \mu_G[d\boldsymbol{\phi} = F_\theta(d\boldsymbol{\varphi})]\left|DF_\theta^{-1}[\boldsymbol{\phi} = F_\theta(\boldsymbol{\varphi})]\right|^{-1}. \tag{C11}$$

Throughout the proof, we set $\theta = 1 - \Delta$ and notice that the final upper bound on the TVD would still hold for $\theta \leq 1 - \Delta$. The change-of-variable formula (C10) directly gives the element of the (diagonal) Jacobian

$$\begin{aligned}\left|\partial_\phi \varphi\right|^{-1} &= \frac{\cos^2(\phi/2) + (1-\Delta)^2\sin^2(\phi/2)}{1-\Delta}\\ &= \frac{1 - \Delta(2-\Delta)\sin^2(\phi/2)}{1-\Delta},\end{aligned} \tag{C12}$$

which attains the minimum when $\sin^2(\phi/2) = 1$ and the maximum when $\sin^2(\phi/2) = 0$. Thus, we have the following bound on the Jacobian for both the passive and active cases:

$$1 - \Delta \leq \left|\partial_\phi \varphi\right|^{-1} \leq \frac{1}{1-\Delta}, \tag{C13}$$

$$(1-\Delta)^d \leq \left|DF_\theta^{-1}[\boldsymbol{\phi} = F_\theta(\boldsymbol{\varphi})]\right|^{-1} \leq \frac{1}{(1-\Delta)^d}. \tag{C14}$$

At last, to bound the TVD, we express the measures μ_G and μ_G^θ in the same coordinates $\boldsymbol{\varphi}$. For the case of passive FLO, this can be done by directly applying the inverse of the deformation map (46) to each group element $e^{i\phi_j}, j \in [d]$

$$F_{1-\Delta}^{-1}(e^{i\varphi_j}) = \frac{\Delta + (\Delta-2)e^{i\varphi_j}}{\Delta(e^{i\varphi_j}+1) - 2}. \tag{C15}$$

As a result,

$$\begin{aligned}\left|e^{i\phi_k} - e^{i\phi_j}\right| &= \left|F_{1-\Delta}^{-1}(e^{i\varphi_k}) - F_{1-\Delta}^{-1}(e^{i\varphi_j})\right|\\ &= \frac{(1-\Delta)\left|e^{i\varphi_k} - e^{i\varphi_j}\right|}{\left|1 - \frac{\Delta}{2}(e^{i\varphi_j}+1)\right|\left|1 - \frac{\Delta}{2}(e^{i\varphi_k}+1)\right|}\\ &=: \Gamma_{\text{pas}}\left|e^{i\varphi_k} - e^{i\varphi_j}\right|,\end{aligned} \tag{C16}$$

which implies, via Weyl's formula (C3) that the two measures are proportional:

$$\mu_{\text{U}(d)}(\boldsymbol{\varphi}) = \Gamma_{\text{pas}}^{d(d-1)/2}\mu_{\text{U}(d)}^\theta(\boldsymbol{\varphi}). \tag{C17}$$

The proportionality constant Γ_{pas} attains the maximum value when $e^{i\varphi_j} = e^{i\varphi_k} = 1$ and the minimum value when $e^{i\varphi_j} = e^{i\varphi_k} = -1$, giving the following bound:

$$(1-\Delta)^2 \leq \Gamma_{\text{pas}} \leq \frac{1}{(1-\Delta)^2}, \tag{C18}$$

which leads to the bound of the TVD stated in the lemma:

$$\left\|\mu_{\text{U}(d)} - \mu_{\text{U}(d)}^\theta\right\|_{\text{TVD}} \leq \frac{1}{2}\left|1 - (1-\Delta)^{d^2}\right| \leq \frac{d^2\Delta}{2}; \tag{C19}$$

the inequality in the last line can be proved by induction on $d^2 \geq 1$.

Turning to the case of active FLO, the change-of-variable formula (C10) implies that for any $j, k \in [d]$,

$$\begin{aligned}\cos\phi_k - \cos\phi_j &= \frac{(1-\Delta)^2 - \tan^2(\varphi_k/2)}{(1-\Delta)^2 + \tan^2(\varphi_k/2)} - \frac{(1-\Delta)^2 - \tan^2(\varphi_j/2)}{(1-\Delta)^2 + \tan^2(\varphi_j/2)} && \text{(C20)}\\ &= \frac{(1-\Delta)^2(\cos\varphi_k - \cos\varphi_j)}{\left[1 - \Delta(1-\frac{\Delta}{2})(1+\cos\varphi_j)\right]\left[1 - \Delta(1-\frac{\Delta}{2})(1+\cos\varphi_k)\right]} && \text{(C21)}\\ &=: \Gamma_{\text{act}}(\cos\varphi_k - \cos\varphi_j), && \text{(C22)}\end{aligned}$$

where we use $\cos(2\theta) = (1 - \tan^2\theta)(1 + \tan^2\theta)^{-1}$ in the first line and $\tan^2\theta = [1 - \cos(2\theta)][1 + \cos(2\theta)]^{-1}$ in the second line. Thus we have from Weyl's formula (C3) that

$$\mu_{\mathrm{SO}(2d)}(\boldsymbol{\varphi}) = \Gamma_{\mathrm{act}}^{d(d-1)/2} \mu^{\theta}_{\mathrm{SO}(2d)}(\boldsymbol{\varphi}). \tag{C23}$$

The proportionality constant Γ_{act} attains the maximum value when $\cos\varphi_j = 1$ and the minimum value when $\cos\varphi_j = -1$, giving the bound

$$(1-\Delta)^2 \leq \Gamma_{\mathrm{act}} \leq \frac{1}{(1-\Delta)^2}. \tag{C24}$$

Therefore, the TVD is bounded in a similar manner to the passive case.

$$\left\|\mu_{\mathrm{SO}(2d)} - \mu^{\theta}_{\mathrm{SO}(2d)}\right\|_{\mathrm{TVD}} \leq \frac{1}{2}\left|1 - (1-\Delta)^{d^2}\right| \leq \frac{d^2\Delta}{2}. \tag{C25}$$

APPENDIX D: POLYNOMIALS ASSOCIATED TO PROBABILITIES IN FLO CIRCUITS

In this section, we give the degrees of matrix polynomials associated to fermionic representations of $G = \mathrm{U}(d)$ and $G = \mathrm{SO}(2d)$. These polynomials, when evaluated on the Cayley path g_θ in the appropriate group [see Eq. (47)], give rise to polynomials and rational functions θ for the outcome probabilities $p_{\mathbf{x}}[\Pi(g_\theta), \Psi_{\mathrm{in}}] = |\langle\mathbf{x}|\,\Pi(g)\,|\Psi_{\mathrm{in}}\rangle|^2$ in our quantum advantage schemes. The explicit form of these polynomials is used in Sec. VIII when discussing worst-to-average-case reductions.

We start with discussing the passive FLO case and then the active FLO case.

It is useful to introduce the following notation. Given a $d \times d$ matrix M and two subsets of indices $\mathcal{X}, \mathcal{Y} \subset [d]$ with cardinality n, where $\mathcal{X} = \{a_1, a_2, \ldots a_n\}$ ($a_i < a_j$ if $i < j$) and $\mathcal{Y} = \{b_1, b_2, \ldots, b_n\}$ ($b_i < b_j$ if $i < j$), we define $M_{\mathcal{X},\mathcal{Y}}$ as the $n \times n$ matrix with entries

$$(M_{\mathcal{X},\mathcal{Y}})_{k,\ell} = M_{a_k, b_\ell},\ k, \ell = 1, \ldots n. \tag{D1}$$

Lemma 10: *Given two Fock basis states $|\mathcal{X}\rangle, |\mathcal{Y}\rangle \in \bigwedge^n(\mathbb{C}^d)$ and a $U \in U(d)$, the the amplitude between $|\mathcal{X}\rangle$ and $\Pi_{\mathrm{pas}}(U)\,|\mathcal{Y}\rangle$ is provided by the expression*

$$\langle\mathcal{X}|\,\Pi_{\mathrm{pas}}(U)\,|\mathcal{Y}\rangle = \det(U_{\mathcal{X},\mathcal{Y}}). \tag{D2}$$

Proof. Let $\mathcal{X} = \{a_1, a_2, \ldots a_n\}$ ($a_i < a_j$ if $i < j$) and $\mathcal{Y} = \{b_1, b_2, \ldots, b_n\}$ ($b_i < b_j$ if $i < j$). By definition we have that

$$\begin{aligned}\Pi_{\mathrm{pas}}(U)\,|\mathcal{Y}\rangle &= U^{\otimes n}\,|b_1\rangle \wedge |b_2\rangle \wedge \cdots \wedge |b_n\rangle \\ &= |\xi_1\rangle \wedge |\xi_2\rangle \wedge \cdots \wedge |\xi_n\rangle\,,\end{aligned} \tag{D3}$$

where

$$|\xi_\ell\rangle = U\,|b_\ell\rangle = \sum_{j=1}^{d} U_{j,b_\ell}\,|j\rangle\,,\ \ell = 1, \ldots, n. \tag{D4}$$

Using the last two equations and Eq. (6), we can deduce that

$$\begin{aligned}&\langle\mathcal{X}|\,\Pi_{\mathrm{pas}}(U)\,|\mathcal{Y}\rangle = \det(C), \\ &C_{k,\ell} = \langle a_k|\xi_\ell\rangle = \langle a_k|\sum_{j=1}^{d} U_{j,b_\ell}\,|j\rangle = U_{a_k,b_\ell} = (U_{\mathcal{X},\mathcal{Y}})_{k,\ell},\end{aligned} \tag{D5}$$

which proves the statement. ■

This lemma allows us to directly obtain the following result.

Proposition 1: (Degrees of polynomials describing probabilities associated to passive FLO circuits.). *Consider a state $|\Psi\rangle \in \bigwedge^n(\mathbb{C}^d)$. For an arbitrary $U \in U(d)$ the outcome probability $p_{\mathbf{x}}[\Pi_{\mathrm{pas}}(U), \Psi] = |\langle\mathbf{x}|\,\Pi_{\mathrm{pas}}(U)\,|\Psi\rangle|^2$ is a degree $2n$ homogeneous polynomial in the entries of U and $U^\dagger$.*

Proof. One can expand the vector $|\Psi\rangle$ in terms of the Fock basis states belonging to $\bigwedge^n(\mathbb{C}^d)$ as

$$|\Psi\rangle = \sum_{\substack{\mathcal{Y}\subset[d]\\ |\mathcal{Y}|=n}} c_{\mathcal{Y}}\,|\mathcal{Y}\rangle\,. \tag{D6}$$

Let $\mathcal{X} \subset [d]$ denote the set of indices corresponding to $\mathbf{x}$ as an indicator function (i.e., $|\mathbf{x}\rangle = |\mathcal{X}\rangle$). Using Lemma 9, we can write the relevant amplitude as

$$\begin{aligned}\langle\mathbf{x}|\,\Pi_{\mathrm{pas}}(U)\,|\Psi\rangle &= \sum_{\substack{\mathcal{Y}\subset[d]\\ |\mathcal{Y}|=n}} c_{\mathcal{Y}}\,\langle\mathcal{X}|\,\Pi_{\mathrm{pas}}(U)\,|\mathcal{Y}\rangle \\ &= \sum_{\substack{\mathcal{Y}\subset[d]\\ |\mathcal{Y}|=n}} c_{\mathcal{Y}}\det(U_{\mathcal{X},\mathcal{Y}}).\end{aligned} \tag{D7}$$

As each term in the sum is a determinant of a $n \times n$ submatrix of U, this expression gives a homogeneous polynomial of the entries of U of order n. This in turn directly implies that $p_{\mathbf{x}}[\Pi_{\mathrm{pas}}(U), \Psi] = |\langle\mathbf{x}|\,\Pi_{\mathrm{pas}}(U)\,|\Psi\rangle|^2$ is a degree $2n$ polynomial in the entries of U and $U^\dagger$. ■

Lemma 11: (Polynomial for output amplitude of passive FLO [56]). *Consider the input state* $|\Psi_{\text{in}}\rangle = |\Psi_4\rangle^{\otimes N} \in \bigwedge^{2N}(\mathbb{C}^{4N})$. *For an arbitrary* $U \in U(4N)$ *the outcome amplitude is given by*

$$\langle \mathbf{x}|\,\Pi_{\text{pas}}(U)\,|\Psi\rangle = \frac{1}{\sqrt{2^N}} \sum_{(y_1,\ldots,y_N)\in\{0,1\}^N} \det(U^T_{\{2y_1+1,2y_1+2,\ldots,2y_N+4N-3,2y_N+4N-2\},\mathcal{X}}), \tag{D8}$$

where $U^T_{\{2y_1+1,2y_1+2,\ldots,2y_N+4N-3,2y_N+4N-2\},\mathcal{X}}$ *indicates the transpose of U with the rows not indexed by* $\{2y_1+1, 2y_1+2,\ldots,2y_N+4N-3, 2y_N+4N-2\}$ *and columns not indexed by* $\mathcal{X}$. *Note that this is a degree N polynomial in the entries of U.*

Proof. To derive this polynomial, we rewrite the input fermionic magic state $|\Psi_{\text{in}}\rangle$ as in Eq. (E18).

$$|\Psi_{\text{in}}\rangle = \frac{1}{\sqrt{2^N}} \sum_{\mathcal{Y}\in\mathcal{C}_{\text{in}}} |\mathcal{Y}\rangle \tag{D9}$$

$$= \frac{1}{\sqrt{2^N}} \sum_{\mathbf{y}\in\{0,1\}^N} |\mathcal{Y}_{\mathbf{y}}\rangle, \tag{D10}$$

where $\mathcal{C}_{\text{in}}$ consists of subsets labeled by bitstrings [for more detail, see paragraph after Eq. (E18)]. Using the expression for the output amplitude in Proposition 1 we write

$$\langle \mathbf{x}|\,\Pi_{\text{pas}}(U)\,|\Psi\rangle$$

$$= \frac{1}{\sqrt{2^N}} \sum_{\mathbf{y}\in\{0,1\}^N} \det(U_{\mathcal{X},\mathcal{Y}_{\mathbf{y}}}) \tag{D11}$$

$$= \frac{1}{\sqrt{2^N}} \sum_{(y_1,\ldots,y_N)\in\{0,1\}^N} \times \det(U^T_{\{2y_1+1,2y_1+2,\ldots,2y_N+4N-3,2y_N+4N-2\},\mathcal{X}}), \tag{D12}$$

where in the last line we replace the definition of $\mathcal{Y}_y$ and also use the fact that the determinant is invariant under the transpose. ■

The expression in Eq. (D8) can be rewritten as a mixed discriminant

$$D_{2,2}(v_1,\ldots,v_{4N}) = \frac{1}{\sqrt{2^N}} \sum_{\substack{i_k=0,1,\\ k=1,\ldots,N}} \det \begin{bmatrix} v_{2i_1+1} \\ v_{2i_1+2} \\ v_{2i_2+5} \\ v_{2i_2+6} \\ \vdots \\ v_{2i_N+4N-3} \\ v_{2i_N+4N-2} \end{bmatrix}, \tag{D13}$$

here v_k correspond to the rows of the matrix U^T in Eq. (D12) with the columns not indexed by $\mathbf{x}$ removed. This polynomial over entries of matrices of size $2N \times N$ is found to be $\#P$-hard in the general case [56]. This is proven by reducing the computation of the permanent of a weighted adjacency matrix to these polynomials of a transformed adjacency matrix with polynomial overhead.

Remark 21: For the hardness of sampling what is actually required is the $\#P$-hardness of computing the square of the amplitude. In Ref. [56] the permanents used involved only positive numbers and thus there is no issue in establishing $\#P$-hardness for the probabilities.

Next we turn to studying the output probabilities after an active FLO evolution. It is useful to introduce the following notation: given a set of (Majorana) indices $\mathcal{A} = \{a_1, a_2, \ldots a_k\} \subset [2d]$ (with $a_i < a_j$ if $i < j$), we define

$$m_{\mathcal{A}} = m_{a_1} m_{a_2} \cdots m_{a_k}. \tag{D14}$$

These Majorana monomials define an orthogonal (but not orthonormal) basis in the space of operators with respect to the Hilbert-Schmidt scalar product

$$\operatorname{tr}(m_{\mathcal{A}} m^\dagger_{\mathcal{B}}) = (-1)^{f(|\mathcal{B}|)} \operatorname{tr}(m_{\mathcal{A}} m_{\mathcal{B}}) = \delta_{\mathcal{A},\mathcal{B}} \frac{1}{2^d}, \tag{D15}$$

where $f(n) = 1$ if $(n \mod 4) \in \{2,3\}$ and $f(n) = 0$ otherwise.

Consider a subset $\mathcal{A} = a_1, a_2 \ldots a_k \subset [2d]$ (with $a_i < a_j$ if $i < j$), then from Eq. (11) and the Majorana anticommutation relations it follows for $O \in \mathrm{SO}(2d)$ that

$$\Pi_{\text{act}}(O) m_{\mathcal{A}} \Pi_{\text{act}}(O)^\dagger$$

$$= \sum_{b_1,\ldots b_k=1}^{d} \epsilon_{b_1,b_2,\ldots,b_k} O_{a_1,b_1} O_{a_2,b_2} \cdots O_{a_k,b_k} m_{\{b_1,\ldots b_k\}}. \tag{D16}$$

Proposition 2: (Degrees of polynomials describing probabilities associated to active FLO circuits.) *Consider a state* $|\Psi\rangle \in$. *For an arbitrary* $O \in SO(2d)$ *the outcome probability* $p_{\mathbf{x}}[\Pi_{\text{act}}(O), \Psi] = |\langle \mathbf{x}|\,\Pi_{\text{act}}(O)\,|\Psi\rangle|^2$ *is a degree d polynomial in the entries of O.*

Proof. Let us consider the expansion of $|\mathbf{x}\rangle\langle\mathbf{x}|$ and Ψ in terms of Majorana monomials

$$|\mathbf{x}\rangle\langle\mathbf{x}| = \sum_{\mathcal{A}\subset[d]} (a_{\mathcal{A}}\, m_{\mathcal{A}} + b_{\mathcal{A}}\, Q\, m_{\mathcal{A}}),$$
$$\Psi = \sum_{\mathcal{B}\subset[d]} (c_{\mathcal{B}}\, m_{\mathcal{B}} + d_{\mathcal{B}}\, Q\, m_{\mathcal{B}}). \tag{D17}$$

Using this and Eqs. (D15) and (D16) can now write the outcome probability as

$$\begin{aligned} p_{\mathbf{x}}[\Pi_{\text{act}}(O),\Psi] &= \text{tr}\left[|\mathbf{x}\rangle\langle\mathbf{x}|\;\Pi_{\text{act}}(O)\Psi\Pi_{\text{act}}(O)^{\dagger}\right] \\ &= \sum_{\mathcal{A},\mathcal{B}\subset[d]} \{a_{\mathcal{A}}c_{\mathcal{B}}\,\text{tr}\left[m_{\mathcal{A}}\Pi_{\text{act}}(O)m_{\mathcal{B}}\Pi_{\text{act}}(O)^{\dagger}\right] + b_{\mathcal{A}}d_{\mathcal{B}}\,\text{tr}\left[Qm_{\mathcal{A}}\Pi_{\text{act}}(O)Qm_{\mathcal{B}}\Pi_{\text{act}}(O)^{\dagger}\right]\} \\ &= \sum_{k=0}^{d} \sum_{\substack{\mathcal{A},\mathcal{B}\subset[d] \\ |\mathcal{A}|=|\mathcal{B}|=k}} w_{\mathcal{A},\mathcal{B}} \sum_{\ell_1,\dots\ell_k=1}^{d} \epsilon_{\ell_1,\ell_2,\dots,\ell_k}\delta_{\mathcal{A},\{\ell_1,\dots,\ell_k\}}\, O_{b_1,\ell_1}O_{b_2,\ell_2}\cdots O_{b_k,\ell_k}, \end{aligned} \tag{D18}$$

where $w_{\mathcal{A},\mathcal{B}} = (-1)^{f(|\mathcal{A}|)}/2^d[a_{\mathcal{A}}c_{\mathcal{B}} + (-1)^k b_{\mathcal{A}}d_{\mathcal{B}}]$. Since each term in the sum is a degree d or less polynomial in the entries of O the theorem is proved. ■

Definition 5: (Degree of rational functions). Let $P(\theta)$, $Q(\theta)$ be polynomials of degree d_1 and d_2, respectively. Let $R(\theta) = P(\theta)/Q(\theta)$ be the corresponding rational function. Assume that P and Q do not have nonconstant polynomial divisors. Then, we define the rational degree of R as the pair $\deg(R) = (d_1, d_2)$

The following results state that FLO circuit representations of elements of the appropriate symmetry group G, when evaluated on Cayley paths, give rise to outcome probabilities that are rational functions of low degree (in number of modes d and number of particles n).

Lemma 12: (Degrees of rational functions describing probabilities associated to interpolation of FLO circuits.) *Let G be equal to $U(d)$ or $SO(2d)$. Let $g_0, g \in G$ be a fixed elements of the group G. Consider a rational path in the group defined by interpolation via Cayley path*

$$g_\theta = g_0 F_\theta(g),\ \theta \in [0,1]. \tag{D19}$$

Let now $\Pi : G \to U(\mathcal{H})$ be the appropriate representation of G describing appropriate class of FLO circuits [$G = U(d)$, $\Pi = \Pi_{\text{pas}}$, $\mathcal{H} = \bigwedge^n(\mathbb{C}^d)$ for passive FLO and $G = SO(2d)$, $\Pi = \Pi_{\text{act}}$, $\mathcal{H} = \mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^d)$ for active FLO]. Let us fix $|\Psi\rangle \in \mathcal{H}$ and a Fock state $|\mathbf{x}\rangle \in \mathcal{H}$. Then the outcome probability

$$R_{g_0,g}(\theta) = \text{tr}[|\mathbf{x}\rangle\langle\mathbf{x}|\Pi(g_\theta)\rho\Pi(g_\theta)^{\dagger}] \tag{D20}$$

viewed as a function of parameter θ is a rational function of degrees.

$$\begin{aligned} &\textit{Passive FLO:} \quad \deg(R_{g_0,g}) = (2dn, 2dn). \\ &\textit{Active FLO:} \quad \deg(R_{g_0,g}) = (2d^2, 2d^2). \end{aligned} \tag{D21}$$

Moreover, the denominator of the rational functions are given by the following.

$$\begin{aligned} &\textit{Passive FLO:} \quad Q_g(\theta) = \prod_{j=1}^{d}[1 + \theta^2\tan^2(\phi_j/2)]^n. \\ &\textit{Active FLO:} \quad Q_g(\theta) = \prod_{j=1}^{d}[1 + \theta^2\tan^2(\phi_j/2)]^d, \end{aligned} \tag{D22}$$

where $\phi_j, j \in [d]$ are phases of generalized eigenvalues of matrix g belonging to the suitable group G and thus $Q_g(\theta)$ can be efficiently computed (see Sec. VII).

Proof. We begin by proving the passive FLO case. Recall from Eq. (55) that g_θ was expressed as a matrix with entries of degree (d,d) on θ. By virtue of Proposition 1, we know that $p_{\mathbf{x}}[\Pi_{\text{pas}}(g_\theta),\Psi] = R_{g_0,g}(\theta)$ is a polynomial of degree $2n$ on the entries of g_θ, which immediately implies the degree on θ is $\deg(R_{g_0,g}) = (2dn, 2dn)$. The denominator of the rational functions in g_θ is given by Eq. (56), from the expression for the amplitude in Proposition 1, we know that the denominator in $R_{g_0,g}$ must be of the form $\left|\mathcal{Q}_g(\theta)^n\right|^2$, which gives the result from $\prod_{j=1}^d[1+\theta^2\tan^2(\phi_j/2)]^n$.

For the active case, we obtain from Eq. (60) that g_θ is a matrix with entries that are polynomials of degree $(2d,2d)$. Then by Proposition 2, $p_{\mathbf{x}}[\Pi_{\rm act}(g_\theta),\Psi]$ is of degree d on the entries of g_θ implying $\deg(R_{g_0,g})=(2d^2,2d^2)$. The denominator $Q_g(\theta)$ is obtained by noting that the expression in Eq. (62) for $\mathcal{Q}_g(\theta)$ appears as the denominator in each entry of g_θ and by Proposition 2 the degree on this denominator is d, thus proving the result. ■

APPENDIX E: DETAILS OF COMPUTATIONS FOR ANTICONCENTRATION

1. Passive FLO

In this part we give detailed computations related to establishing upper bound in Eq. (39) for the case of passive FLO:

$$\operatorname{tr}(\mathbb{P}_{\rm pas}\Psi_{\rm in}\otimes\Psi_{\rm in})\le\frac{C_{\rm pas}}{N},\ \text{for}\ C_{\rm pas}=5.7. \tag{E1}$$

The prove of the above inequality is split into three parts. First, in Lemma 12 we give an explicit form of $\mathbb{P}_{\rm pas}$. Second, in Lemma 13 we find an upper bound on $\operatorname{tr}(\mathbb{P}_{\rm pas}\Psi_{\rm in}\otimes\Psi_{\rm in})$ via combinatorial expression that can be efficiently computed for any fixed value of N. Finally, in Lemma 17 given in Part E 3 of the Appendix we prove an upper bound to the said combinatorial expression, which yields Eq. (E1).

Lemma 13: (Projector for passive fermionic linear optics.) *Let $\bigwedge^n\left(\mathbb{C}^d\right)$ be a fermionic n-particle representation of $U(d)$ $(d\ge n)$. Let $\mathbb{P}_{\rm pas}$ be the projector onto a unique irreducible representation $\tilde{\mathcal{H}}_f\subset\bigwedge^n\left(\mathbb{C}^d\right)\otimes\bigwedge^n\left(\mathbb{C}^d\right)$ of $U(d)$ such that $|\mathbf{n}\rangle\otimes|\mathbf{n}\rangle\in\tilde{\mathcal{H}}_f$, where $|\mathbf{n}\rangle$ is a n-particle Fock state. Then, for any $\rho\in\mathcal{D}[\bigwedge^n(\mathbb{C}^d)]$ we have*

$$\operatorname{tr}(\mathbb{P}_{\rm pas}\rho\otimes\rho)=\frac{1}{n+1}\sum_{k=0}^n\binom{n}{k}\operatorname{tr}(\rho_k^2), \tag{E2}$$

where $\rho_k=\operatorname{tr}_{n-k}(\rho)$ is a k-particle reduction of ρ. Moreover, the dimension of $\tilde{\mathcal{H}}_f$ equals

$$|\tilde{\mathcal{H}}_f|=\binom{d}{n}^2\frac{d+1}{(d-n+1)(n+1)}. \tag{E3}$$

Proof. We consider $\bigwedge^n(\mathbb{C}^d)$ as antisymmetric subspace of the Hilbert space of n distinguishable particles: $\bigwedge^n(\mathbb{C}^d)\subset(\mathbb{C}^d)^{\otimes n}$ with d-dimensional single-particle Hilbert spaces. Therefore, also $\bigwedge^n(\mathbb{C}^d)\otimes\bigwedge^n(\mathbb{C}^d)$ can be considered as a subspace of $2n$ distinguishable particles:

$$\bigwedge^n(\mathbb{C}^d)\otimes\bigwedge^n(\mathbb{C}^d)\subset(\mathbb{C}^d)^{\otimes n}\otimes(\mathbb{C}^d)^{\otimes n}. \tag{E4}$$

Let us now label particles entering the first factor of the latter tensor product by $1,\dots,n$ and by $1',\dots,n'$ particles entering the second factor. In Ref. [101] it was proven that

$$\mathbb{P}_{\rm pas}=\frac{2^n}{n+1}\mathbb{P}^{\{1,\dots,n\}}_{\rm asym}\mathbb{P}^{\{1',\dots,n'\}}_{\rm asym}\left(\prod_{k=1}^n\mathbb{P}^{k,k'}_{\rm sym}\right)\mathbb{P}^{\{1,\dots,n\}}_{\rm asym}\mathbb{P}^{\{1',\dots,n'\}}_{\rm asym}. \tag{E5}$$

In the above $\mathbb{P}^{k,k'}_{\rm sym}=\frac{1}{2}(\mathbb{I}\otimes\mathbb{I}+\mathbb{S}^{k,k'})$ is the projector onto a subspace of $(\mathbb{C}^d)^{\otimes n}\otimes(\mathbb{C}^d)^{\otimes n}$, which is symmetric upon interchange of particles k and k' (by $\mathbb{S}^{k,k'}$ we denote the unitary operator that swaps particles k and k'). Moreover, $\mathbb{P}^{\mathcal{A}}_{\rm asym}$ denotes the projector onto a subspace, which is antisymmetric under exchange of particles in a subset $\mathcal{A}$. Now for $\rho\in\mathcal{D}\left[\bigwedge^n(\mathbb{C}^d)\right]$ we have

$$\mathbb{P}^{\{1,\dots,n\}}_{\rm asym}\mathbb{P}^{\{1',\dots,n'\}}_{\rm asym}\rho\otimes\rho=\rho\otimes\rho \tag{E6}$$

and therefore

$$\operatorname{tr}(\mathbb{P}_{\rm pas}\rho\otimes\rho)=\frac{2^n}{n+1}\operatorname{tr}\left[\left(\prod_{k=1}^n\mathbb{P}^{k,k'}_{sym}\right)\rho\otimes\rho\right]. \tag{E7}$$

Using the definition of $\mathbb{P}^{k,k'}_{sym}$ we get the expansion

$$\prod_{k=1}^n\mathbb{P}^{k,k'}_{sym}=\frac{1}{2^n}\sum_{\mathcal{X}\subset[d]}\prod_{i\in\mathcal{X}}\mathbb{S}^{i,i'}, \tag{E8}$$

where the summation is over subsets X of $[d]=\{1,\dots,d\}$. Using a well-known connection between partial swaps and purities of reduced density matrices (see, for example, Ref. [115]):

$$\operatorname{tr}\left(\prod_{i\in\mathcal{X}}\mathbb{S}^{i,i'}\rho\otimes\rho\right)=\operatorname{tr}(\rho_X^2), \tag{E9}$$

where $\rho_X=\operatorname{tr}_{[d]\setminus\mathcal{X}}(\rho)$ is the reduction of ρ to particles in X. From the symmetry of ρ we have $\operatorname{tr}(\rho_X^2)=\operatorname{tr}(\rho_k^2)$, where $k=|X|$ (size of the set X). Inserting this into (E8) and (E7) we finally obtain

$$\operatorname{tr}(\mathbb{P}_{\rm pas}\rho\otimes\rho)=\frac{1}{n+1}\sum_{k=0}^n\binom{n}{k}\operatorname{tr}(\rho_k^2). \tag{E10}$$

The formula for the dimension, Eq. (E3), follows from the fact that the Hilbert space $\tilde{\mathcal{H}}_f$ is a carrier space of an irreducible representation of U(d) labeled by a Young diagram having two columns each of which has n rows. Formulas for dimensions of such irreducible representations are known (see, for example, Ref. [116]) and were used previously in the context of detection of mixed states that cannot be decomposed as a convex combination of Slater determinants [102]. ■

Note that for all n-particle pure states we have $\mathrm{tr}(\rho_k^2) = \mathrm{tr}(\rho_{n-k}^2)$. This observation gives us the following.

Corollary 2: *Let $\Psi \in \mathcal{D}[\bigwedge^{2m}(\mathbb{C}^d)]$ be a pure state. Let $\mathbb{P}_{\mathrm{pas}}$ be defined as in Lemma 12. We then have*

$$\mathrm{tr}(\mathbb{P}_{\mathrm{pas}}\Psi \otimes \Psi) = \frac{1}{2m+1}\left[2\sum_{k=0}^{m-1}\binom{2m}{k}\mathrm{tr}(\rho_k^2) + \binom{2m}{m}\mathrm{tr}(\rho_m^2)\right]. \tag{E11}$$

We now proceed with some further technical results, which allows us to compute $\mathrm{tr}(\mathbb{P}_{\mathrm{pas}}\rho_{\mathrm{in}} \otimes \rho_{\mathrm{in}})$.

For a set of indices $\mathcal{X} = \{x_1, x_2, \ldots, x_n\} \subset [d]$ where $x_i < x_j$ if $i < j$, and a subset of it $\mathcal{S} = \{x_{\ell_1}, x_{\ell_2}, \ldots, x_{\ell_k}\} \subset \mathcal{X}$ it is useful to introduce the following sign:

$$\begin{aligned}(-1)^{J(\mathcal{X},\mathcal{S})},\ \text{where}\ & J(\mathcal{X},\mathcal{S}) \\ &= \ell_1 + \ell_2 + \ldots + \ell_k + \frac{k(k-1)}{2}.\end{aligned} \tag{E12}$$

This notation allows us to express in a compact way the following matrix element: for any two Fock basis states $|\mathcal{X}\rangle, |\mathcal{Y}\rangle \in \bigwedge^n \mathbb{C}^d$ belonging to the index sets $\mathcal{X}, \mathcal{Y} \subset [d]$, we have that

$$\langle\mathcal{X}| f_{s_1}^\dagger f_{s_2}^\dagger \cdots f_{s_k}^\dagger f_{q_k} \cdots f_{q_2} f_{q_1} |\mathcal{Y}\rangle = \begin{cases}\delta_{\mathcal{X}\setminus\mathcal{S},\mathcal{Y}\setminus\mathcal{Q}}\, \epsilon_{s_1,\ldots,s_k}\epsilon_{q_1,\ldots,q_k}(-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{Q})}\ \text{if}\ \mathcal{S}\subset\mathcal{X}\ \text{and}\ \mathcal{Q}\subset\mathcal{Y}, \\ 0\ \text{else},\end{cases} \tag{E13}$$

where $\mathcal{S} = \{s_1, s_2 \ldots s_k\}$ and $\mathcal{Q} = \{q_1, q_2, \ldots q_k\}$.

Proposition 3: *Let $|\mathcal{X}\rangle, |\mathcal{Y}\rangle \in \bigwedge^n(\mathbb{C}^d)$ be fermionic n-particle Fock states corresponding to n-element subsets $\mathcal{X}, \mathcal{Y} \subset [d]$ (cf. notation introduced in Sec. II), then for any $k = 0, \ldots, n$ we have*

$$\mathrm{tr}_k(|\mathcal{X}\rangle\langle Y|) = \frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})}\, |\mathcal{X}\setminus\mathcal{S}\rangle\langle\mathcal{Y}\setminus\mathcal{S}|\,. \tag{E14}$$

Note that the notation used in the above expression implies $\mathrm{tr}_k(|\mathcal{X}\rangle\langle\mathcal{Y}|) = 0$ *if* $|\mathcal{X}\cap\mathcal{Y}| < k$.

Proof. For any two states $|\Psi\rangle, |\Phi\rangle \in \bigwedge^n \mathbb{C}^d \subset (\mathbb{C}^d)^{\otimes n}$, the k-fold partial trace (with regards to the tensor product structure) results in an operator $O = \mathrm{tr}_k(|\Psi\rangle\langle\Phi|) \in \mathcal{B}(\bigwedge^\ell \mathbb{C}^d) \subset \mathcal{B}[(\mathbb{C}^d)^{\otimes\ell}]$ (with $\ell = n - k$) that has the following matrix elements [117]:

$$\langle v_1| \otimes \langle v_2| \otimes \cdots \langle v_\ell|\ O\ |w_1\rangle \otimes |w_2\rangle \otimes \cdots \otimes |w_\ell\rangle = \frac{1}{\binom{n}{k}} \langle\Phi|\, (f_1^\dagger)^{v_1}(f_2^\dagger)^{v_2} \cdots (f_\ell^\dagger)^{v_\ell}(f_\ell)^{w_\ell} \cdots (f_2)^{w_2}(f_1)^{w_1}\, |\Psi\rangle\ . \tag{E15}$$

Inserting in this equation the $|\Psi\rangle = |\mathcal{X}\rangle$ and $|\Phi\rangle = |\mathcal{Y}\rangle$ and using Eq. (E13), we get that

$$\begin{aligned}&\langle v_1| \otimes \langle v_2| \otimes \cdots \langle v_\ell|\ O\ |w_1\rangle \otimes |w_2\rangle \otimes \cdots \otimes |w_\ell\rangle \\ &= \begin{cases}\binom{n}{k}^{-1}\, \delta_{\mathcal{Y}\setminus\mathcal{A},\mathcal{X}\setminus\mathcal{B}}\, \epsilon_{v_1,\ldots,v_k}\epsilon_{w_1,\ldots,w_k}(-1)^{J(\mathcal{Y},\mathcal{A})+J(\mathcal{X},\mathcal{B})}\ \text{if}\ \mathcal{A}\subset\mathcal{Y}\ \text{and}\ \mathcal{B}\subset\mathcal{X}, \\ 0\ \text{else},\end{cases}\end{aligned} \tag{E16}$$

■

where $\mathbf{v} = (v_1, \ldots, v_\ell)$ and $\mathbf{w} = (w_1, \ldots, w_\ell)$ are the indicator bitstrings of the sets $\mathcal{A}$ and $\mathcal{B}$, respectively. Now considering also the following matrix entries:

$$
\begin{aligned}
&\langle v_1| \otimes \langle v_2| \otimes \cdots \langle v_\ell| \left(\frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})} |\mathcal{X}\setminus\mathcal{S}\rangle\langle\mathcal{Y}\setminus\mathcal{S}| \right) |w_1\rangle \otimes |w_2\rangle \otimes \cdots \otimes |w_\ell\rangle \\
&= \operatorname{tr}\left(\frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})} |\mathcal{X}\setminus\mathcal{S}\rangle\langle\mathcal{Y}\setminus\mathcal{S}| (f_1{}^\dagger)^{v_1} \cdots (f_\ell{}^\dagger)^{v_\ell} (f_\ell)^{w_\ell} \cdots (f_1)^{w_1} \right) \\
&= \frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})} \langle\mathcal{Y}\setminus\mathcal{S}| (f_1{}^\dagger)^{v_1} \cdots (f_\ell{}^\dagger)^{v_\ell} (f_\ell)^{w_\ell} \cdots (f_1)^{w_1} |\mathcal{X}\setminus\mathcal{S}\rangle \\
&= \frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})} \delta_{\mathcal{Y}\setminus\mathcal{S},\mathcal{A}} \delta_{\mathcal{X}\setminus\mathcal{S},\mathcal{B}}\, \epsilon_{v_1,\ldots,v_k} \epsilon_{w_1,\ldots,w_k} \\
&= \frac{1}{\binom{n}{k}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{J(\mathcal{X},\mathcal{S})+J(\mathcal{Y},\mathcal{S})} \delta_{\mathcal{Y}\setminus\mathcal{A},\mathcal{S}} \delta_{\mathcal{Y}\setminus\mathcal{B},\mathcal{S}}\, \epsilon_{v_1,\ldots,v_k} \epsilon_{w_1,\ldots,w_k} \\
&= \begin{cases} \binom{n}{k}^{-1} \delta_{\mathcal{Y}\setminus\mathcal{A},\mathcal{X}\setminus\mathcal{B}}\, \epsilon_{v_1,\ldots,v_k} \epsilon_{w_1,\ldots,w_k} (-1)^{J(\mathcal{Y},\mathcal{A})+J(\mathcal{X},\mathcal{B})} \text{ if } \mathcal{A}\subset\mathcal{Y} \text{ and } \mathcal{B}\subset\mathcal{X}, \\ 0, \text{ else}, \end{cases}
\end{aligned}
\tag{E17}
$$

where we use that $(-1)^{J(\mathcal{Y},\mathcal{A})+J(\mathcal{X},\mathcal{B})} = (-1)^{J(\mathcal{Y},\mathcal{S})+J(\mathcal{X},\mathcal{S})}$, which follows from the fact that $(-1)^{J(\mathcal{X},\mathcal{S})} = (-1)^{J(\mathcal{X},\mathcal{X}\setminus\mathcal{S})+|\mathcal{X}|\cdot|\mathcal{S}|}$. Thus, the matrix elements of Eqs. (E16) and (E17) coincide, which proves the propositions. ■

We introduce the convenient notation for $|\Psi_{\text{in}}\rangle$:

$$|\Psi_{\text{in}}\rangle = \frac{1}{\sqrt{2^N}} \sum_{\mathcal{X}\in\mathcal{C}_{\text{in}}} |\mathcal{X}\rangle\,, \tag{E18}$$

where $\mathcal{C}_{\text{in}}$ is a collection of subsets of $[4N]$ that appear in the decomposition of $|\Psi_{\text{in}}\rangle$. Note that from the definition of $|\Psi_{\text{in}}\rangle$ it follows that subsets are labeled by bitstrings $\mathbf{x} = (x_1, \ldots, x_N)$, where $x_i \in \{0, 1\}$ labels which pair of the neighboring physical modes are occupied in a given quadropule of modes. For $N = 2$ we have four possible subsets belonging to $\mathcal{C}_{\text{in}}$

$$
\begin{aligned}
&\mathcal{X}_{00} = \{1, 2, 5, 6\},\ \mathcal{X}_{01} = \{1, 2, 7, 8\},\ \mathcal{X}_{10} = \{3, 4, 5, 6\}, \\
&\mathcal{X}_{11} = \{3, 4, 7, 8\}.
\end{aligned}
\tag{E19}
$$

For general N the collection $\mathcal{C}_{\text{in}}$ consists of the following subsets labeled by bitstrings $\mathbf{x}$

$$
\begin{aligned}
\mathcal{X}_{\mathbf{x}} = \{&1 + 2x_1, 2 + 2x_1, 5 + 2x_2, 6 + 2x_2, \ldots, 4i - 3 \\
&+ 2x_i, 4i - 2 + 2x_i, \ldots, 4N - 3 \\
&+ 2x_N, 4N - 2 + 2x_N\}.
\end{aligned}
\tag{E20}
$$

The formula from Lemma 3 allows us to obtain bounds for the purities of reduced density matrices of Ψ_{in}.

Proposition 4: (Bounds on purities of reduced density matrices of ρ_{in}.) *Consider the setting of this paper, i.e., $d = 4N$ and $n = 2N$, where N is the number of quadruples used in our quantum advantage proposal. Let $\Psi_{\text{in}} \in \mathcal{D}(\mathcal{H}_f)$ be the input state. Then, for $k = 0, \ldots, N$ we have*

$$\operatorname{tr}\left[\operatorname{tr}_k(\Psi_{\text{in}})^2\right] \le \frac{1}{\binom{2N}{k}^2} \sum_{l=0}^{\lfloor k/2 \rfloor} \frac{N!}{l!(k-2l)!(N-k+l)!}. \tag{E21}$$

Proof. We use the decomposition of the state vector $|\Psi_{\text{in}}\rangle$ given in Eq. (E18) and obtain

$$\Psi_{\text{in}} = \frac{1}{2^N} \sum_{\mathcal{X},\mathcal{Y}\in\mathcal{C}_{\text{in}}} |\mathcal{X}\rangle\langle\mathcal{Y}|\,. \tag{E22}$$

Employing (E14) and denoting $J(\mathcal{X},\mathcal{S}) + J(\mathcal{Y},\mathcal{S}) = K(\mathcal{X},\mathcal{Y},\mathcal{S})$ we obtain (remember that $n = 2N$)

$$
\begin{aligned}
\operatorname{tr}_k(\Psi_{\text{in}}) = \frac{1}{2^N\binom{2N}{k}} &\sum_{\mathcal{X},\mathcal{Y}\in\mathcal{C}_{\text{in}}} \sum_{\mathcal{S}\in\binom{\mathcal{X}\cap\mathcal{Y}}{k}} (-1)^{K(\mathcal{X},\mathcal{Y},\mathcal{S})} \\
&\times |\mathcal{X}\setminus\mathcal{S}\rangle\langle\mathcal{Y}\setminus\mathcal{S}|\,.
\end{aligned}
\tag{E23}
$$

By reordering the sum we obtain

$$\mathrm{tr}_k(\Psi_{\mathrm{in}}) = \frac{1}{2^N\binom{2N}{k}} \sum_{\mathcal{X}',\mathcal{Y}'\in\binom{[4N]}{2N-k}} \left|\mathcal{X}'\right\rangle\!\left\langle\mathcal{Y}'\right| \times \sum_{\substack{\mathcal{S}\in\binom{[4N]}{k},\ \mathcal{X},\mathcal{Y}\in\mathcal{C}_{\mathrm{in}} \\ \text{s.t. } \mathcal{X}\setminus\mathcal{S}=\mathcal{X}',\mathcal{Y}\setminus\mathcal{S}=\mathcal{X}'}} (-1)^{K(\mathcal{X},\mathcal{Y},\mathcal{S})}, \quad \text{(E24)}$$

where the second sum is a combinatorial term that gives a coefficient, and a particular operator $\left|\mathcal{X}'\right\rangle\!\left\langle\mathcal{Y}'\right|$ appears. Crucially, operators $\left|\mathcal{X}'\right\rangle\!\left\langle\mathcal{Y}'\right|$, $|\mathcal{X}'| = |\mathcal{Y}'| = 2N-k$ are orthonormal with respect to the Hilbert-Schmidt inner product in $\mathcal{B}[\bigwedge^{2N-k}(\mathbb{C}^{4N})]$. Therefore, in order to bound purity of $\mathrm{tr}_k(\rho_{\mathrm{in}})$ it suffices to count the number of terms in the second sum in Eq. (E24):

$$\mathrm{tr}\left[\mathrm{tr}_k(\Psi_{\mathrm{in}})^2\right] \leq \frac{1}{2^{2N}\binom{2N}{k}^2} \sum_{\mathcal{X}',\mathcal{Y}'\in\binom{[4N]}{2N-k}} \mathcal{N}(\mathcal{X}',\mathcal{Y}')^2, \quad \text{(E25)}$$

where

$$\mathcal{N}(\mathcal{X}',\mathcal{Y}') = \left|\left\{(\mathcal{S},\mathcal{X},\mathcal{Y}) \mid \mathcal{S}\in\binom{[4N]}{k},\ \mathcal{X},\mathcal{Y}\in\mathcal{C}_{\mathrm{in}},\ \mathcal{X}\setminus\mathcal{S}=\mathcal{X}',\ \mathcal{Y}\setminus\mathcal{S}=\mathcal{X}'\right\}\right|. \quad \text{(E26)}$$

In what follows, in order to make our considerations less abstract, we refer to elements of subsets involved as "particles." To compute $N(\mathcal{X}',\mathcal{Y}')$ we note that $\mathcal{X}',\mathcal{Y}'$ for which $\mathcal{N}(\mathcal{X}',\mathcal{Y}')\neq 0$ must arise from subtracting from $\mathcal{X}\in\mathcal{C}_{\mathrm{in}}$ particles occupying subset $\mathcal{S}$. Since particles corresponding to $\mathcal{X}\in\mathcal{C}_{\mathrm{in}}$ occupy only two out of four possible modes in every quadropule of modes in a "binary fashion" [see Eq. (E19)], This imposes constraints on the possible configurations of particles from $\mathcal{X}'$ in every quadropole. Specifically, consider the quadropule of physical modes $\mathcal{A}=\{1,2,3,4\}$. Let $\mathcal{X}'_{\mathcal{A}}=\mathcal{X}'\cap\mathcal{A}$. We have seven possibilities for the set $\mathcal{X}'_{\mathcal{A}}$:

$$\mathcal{X}'^{\mathrm{N1}}_{\mathcal{A}} = \{1,2\},\ \mathcal{X}'^{\mathrm{N2}}_{\mathcal{A}} = \{3,4\},\ \mathcal{X}'^{\mathrm{B}}_{\mathcal{A}} = \emptyset, \quad \text{(E27)}$$

$$\mathcal{X}'^{\mathrm{F1}}_{\mathcal{A}} = \{1\},\ \mathcal{X}'^{\mathrm{F2}}_{\mathcal{A}} = \{2\},\ \mathcal{X}'^{\mathrm{F3}}_{\mathcal{A}} = \{3\},\ \mathcal{X}'^{\mathrm{F4}}_{\mathcal{A}} = \{4\}. \quad \text{(E28)}$$

All other forms of $\mathcal{X}\cap\mathcal{A}$ yield $\mathcal{N}(\mathcal{X}',\mathcal{Y}')=0$. Under the condition that $\mathcal{X}'$ originates from $\mathcal{X}\in\mathcal{C}_{\mathrm{in}}$ these configurations impose conditions on possible arrangement of lost particles in quadruple $\mathcal{A}$, denoted by $\mathcal{S}_{\mathcal{A}}=\mathcal{S}\cap\mathcal{A}$:

$$\mathcal{S}_{\mathcal{A}}(N1) = \mathcal{S}_{\mathcal{A}}(N2) = \emptyset,\ \mathcal{S}_{\mathcal{A}}(B) = \{1,2\} \text{ or } \mathcal{S}_{\mathcal{A}}(B) = \{3,4\}, \quad \text{(E29)}$$

$$\mathcal{S}_{\mathcal{A}}(F1) = \{2\},\ \mathcal{S}_{\mathcal{A}}(F2) = \{1\},\ \mathcal{S}_{\mathcal{A}}(F3) = \{4\},\ \mathcal{S}_{\mathcal{A}}(F4) = \{3\}. \quad \text{(E30)}$$

This motivates us to introduce *types* of quadruples of $\mathcal{X}'$ whose names are motivated by types of constraints the impose on $\mathcal{S}\cap\mathcal{A}$:

$$T_{\mathcal{A}}(\mathcal{X}') = \begin{cases} \text{NULL} & \text{iff } \mathcal{X}'_{\mathcal{A}} = \{1,2\} or \mathcal{X}'_{\mathcal{A}} = \{3,4\} \\ \text{BINARY} & \text{iff } \mathcal{X}'_{\mathcal{A}} = \emptyset \\ \text{FIXED} & \text{iff } \mathcal{X}'_{\mathcal{A}} \in \{\{1\},\{2\},\{3\},\{4\}\}. \end{cases} \quad \text{(E31)}$$

We repeat the same procedure for other quadruples $\{5,6,7,8\}$, $\{9,10,11,12\}$, etc. To a given $\mathcal{X}'$ we then associate "pattern of types":

$$\mathcal{X}' \longmapsto \mathcal{L}(\mathcal{X}') = \left(l_N[\mathcal{X}'],\ l_B[\mathcal{X}'],\ l_F[\mathcal{X}']\right), \quad \text{(E32)}$$

which lists the number of quadruples of different types in $\mathcal{X}'$. This pattern gives us the number of k-element subsets $\mathcal{S}\in\binom{[4N]}{k}$ contributing to $\mathcal{N}(\mathcal{X}',\mathcal{Y}')$ [cf. Eq. (E26)]. From the considerations given previously $\mathcal{N}_{\mathcal{S}}(\mathcal{X}') = 2^{l_B[\mathcal{X}']}$, where different $\mathcal{S}$ contribute. Let us chose $\mathcal{Y}'$ that is compatible with the pattern of lost particles in $\mathcal{X}'$ is the sense that $\mathcal{X}'\cap\mathcal{Y}'=\emptyset$ and $\mathcal{Y}'$ follows the general constrains of occupations in each quadruples described previously [like the ones stated in Eqs. (E27) and (E28)]. Since for fixed $\mathcal{X}',\mathcal{Y}'$ subset $\mathcal{S}$ uniquely specifies $\mathcal{X},\mathcal{Y}\in\mathcal{C}_{\mathrm{in}}$, we finally get

$$\mathcal{N}(\mathcal{X}',\mathcal{Y}') = 2^{l_B[\mathcal{X}']}. \quad \text{(E33)}$$

It is now easy to see that, for $\mathcal{X}'$ characterized by particular $\mathcal{L}(\mathcal{X}')$, there are exactly

$$\mathcal{N}_{\mathrm{comp}}(\mathcal{X}') = 2^{l_N[\mathcal{X}']} \quad \text{(E34)}$$

different compatible sets $\mathcal{Y}'$. In fact, compatible $\mathcal{Y}'$ necessarily satisfy $\mathcal{L}(\mathcal{Y}')=\mathcal{L}(\mathcal{X}')$. Finally, simple counting argument shows that there are

$$\mathcal{N}(\mathcal{L}) = 2^{2l_F+l_N}\frac{N!}{l_N!l_B!l_F!} \tag{E35}$$

different subsets $\mathcal{X}'$ that have the "pattern type" $\mathcal{L} = (l_N, l_B, l_F)$. Hence the contribution in from subsets $\mathcal{X}', \mathcal{Y}'$ of "pattern type" $\mathcal{L} = (l_N, l_B, l_F)$ to the sum in Eq. (E25) is equal to

$$\mathcal{N}(\mathcal{X}', \mathcal{Y}')^2\mathcal{N}_{\mathrm{comp}}(\mathcal{X}')\mathcal{N}(\mathcal{L}) = 4^{l_N+l_B+l_F}\frac{N!}{l_N!l_B!l_F!} \\ = 2^{2N}\frac{N!}{l_N!l_B!l_F!}. \tag{E36}$$

Parameters l_N, l_B, l_F are not independent because of the following identities: $N = l_N + l_B + l_F$ (this one we already used implicitly) and $2l_B + l_F = 2N - k$. Choosing l_B as an independent parameter applying the above considerations to Eq. (E25) we finally obtain

$$\mathrm{tr}\left[\mathrm{tr}_k(\Psi_{\mathrm{in}})^2\right] \leq \frac{1}{\binom{2N}{k}^2}\sum_{l_B=0}^{\lfloor\frac{k}{2}\rfloor}\frac{N!}{(N-k+l_B)!(k-2l_B)!l_B!}, \tag{E37}$$

where summation range for l_B comes from its definition as the number of quadropules in $\mathcal{X}'$ that are left without particles. ■

Combining Corollary 2 and Proposition 4 we obtain explicitly the upper bound for the expectation value of the projector $\mathbb{P}_{\mathrm{pas}}$

Lemma 14: *Consider the setting of our quantum advantage proposal, i.e., $d = 4N$ and $n = 2N$. Let $\Psi_{\mathrm{in}} \in \mathcal{D}[\bigwedge^{2N}(\mathbb{C}^{4N})]$. Let $\mathbb{P}_{\mathrm{pas}}$ be defined as in Lemma 12. We then have*

$$\mathrm{tr}\left(\mathbb{P}_{\mathrm{pas}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}}\right) \leq \frac{1}{2N+1} \\ \times\left[2\sum_{k=0}^{N-1}\binom{2N}{k}\mathrm{tr}(\rho_k^2) + \binom{2N}{N}\mathrm{tr}(\rho_N^2)\right], \tag{E38}$$

where

$$\mathrm{tr}(\rho_k^2) = \frac{1}{\binom{2N}{k}^2}\sum_{l=0}^{\lfloor k/2\rfloor}\frac{N!}{l!(k-2l)!(N-k+l)!}. \tag{E39}$$

2. Active FLO

We give here computations related to establishing upper bound in Eq. (39) for the case of active FLO:

$$\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}}) \leq \frac{C_{\mathrm{act}}}{\sqrt{\pi N}}, \text{ for } C_{\mathrm{act}} = 16.2. \tag{E40}$$

Similarly to the case of passive FLO the proof divided into three parts. First, in Lemma 14 we give an explicit form of $\mathbb{P}_{\mathrm{act}}$. Second, in Lemma 15 we find an upper bound on $\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}})$ via combinatorial expression that can be efficiently computed for any fixed value of N. Finally, in Lemma 18 given in Part E 3 of the Appendix we prove an upper bound to the said expression, which yields Eq. (E40).

Recall that by m_i, $i = 1, \ldots, 2d$ we denote the standard Majorana operators in the d-mode fermionic Fock space $\mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$ (cf. Sec. II). The fermionic parity operator is given by $Q = i^d\prod_{i=1}^{2d} m_i$.

Lemma 15: (Projector for active fermionic linear optics.) *Let $\mathcal{H}_{\mathrm{act}} = \mathcal{H}^+_{\mathrm{Fock}}\left(\mathbb{C}^d\right)$ be the positive parity subspace of Fock space corresponding to d fermionic modes. Let $\mathbb{P}_{\mathrm{act}}$ be the projector onto a unique irreducible representation $\tilde{\mathcal{H}}_{\mathrm{act}} \subset \mathcal{H}_{\mathrm{act}}\otimes\mathcal{H}_{\mathrm{act}}$ of $SO(2d)$ such that $|\Phi\rangle\otimes|\Phi\rangle \in \tilde{\mathcal{H}}_{\mathrm{act}}$, where Φ are arbitrary pure positive parity Gaussian states. We then have*

$$\mathbb{P}_{\mathrm{act}} = \mathbb{P}_+{\otimes}\mathbb{P}_+\mathbb{P}_0\mathbb{P}_+{\otimes}\mathbb{P}_+, \tag{E41}$$

where $\mathbb{P}_+ = \frac{1}{2}(\mathbb{I}+Q)$ is the orthogonal projector onto $\mathcal{H}^+_{\mathrm{Fock}}(\mathbb{C}^d) \subset \mathcal{H}_{\mathrm{Fock}}(\mathbb{C}^d)$ and

$$\mathbb{P}_0 = \frac{1}{2^{2d}}\sum_{p=0}^{d} C_p \sum_{\mathcal{X}\in\binom{[2d]}{2p}}\prod_{i\in\mathcal{X}} m_i\otimes m_i. \tag{E42}$$

The numbers C_p satisfy $C_p = (-1)^d C_{d-p}$ and for $p \leq \lfloor d/2\rfloor$ we have

$$C_p = (-1)^p\frac{(2p)!(2d-2p)!}{(d!)^2}\binom{d}{p}. \tag{E43}$$

Moreover, the dimension $\tilde{\mathcal{H}}_{\mathrm{act}}$ equals

$$|\tilde{\mathcal{H}}_{\mathrm{act}}| = \frac{1}{2}\binom{2d}{d}. \tag{E44}$$

Proof. The result follows from the characterization of pure fermionic Gaussian states given in Corollary 1 in Ref. [72], which states that a pure state Ψ is a pure fermionic Gaussian state if and only if $\Lambda|\Psi\rangle\otimes|\Psi\rangle = 0$, where Λ is the operator acting on $\mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^d)\otimes\mathcal{H}^{+}_{\text{Fock}}(\mathbb{C}^d)$ introduced previously by Bravyi in Ref. [65]

$$\Lambda=\sum_{i=1}^{2d} m_i\otimes m_i. \tag{E45}$$

Equivalently, Ψ is the pure fermionic Gaussian state if and only if $\mathbb{P}^{\Lambda}_0\left|\Psi\right\rangle\otimes\left|\Psi\right\rangle=0$, where $\mathbb{P}^{\Lambda}_0$ is the projector onto zero eigenspace of Λ. Since we are interested in Gaussian states having positive parity ($Q\left|\Psi\right\rangle=\left|\Psi\right\rangle$), and operators $Q\otimes\mathbb{I}$, $\mathbb{I}\otimes Q$ commute with Λ, we get the following equivalence:

$$\Psi \text{ is pure positive parity fermionic Gaussian state}$$
$$\iff \mathbb{P}_{+}\otimes\mathbb{P}_{+}\mathbb{P}^{\Lambda}_0\mathbb{P}_{+}\otimes\mathbb{P}_{+}\left|\Psi\right\rangle\otimes\left|\Psi\right\rangle=0. \tag{E46}$$

We now show that $\mathbb{P}^{\Lambda}_0=\mathbb{P}_0$. The operator Λ from Eq. (E45) is a sum of $2d$ commuting hermitian operators $m_i\otimes m_i$, which satisfy $(m_i\otimes m_i)^2=\mathbb{I}\otimes\mathbb{I}$. A one-dimensional projector onto a joint eigenspace of M_i corresponding to eigenvalues μ_i, $i\in[2d]$ reads is given by

$$\mathbb{P}_{\boldsymbol{\mu}}=\frac{1}{2^{2d}}\prod_{i=1}^{2d}\left(\mathbb{I}\otimes\mathbb{I}+\mu_i m_i\otimes m_i\right), \tag{E47}$$

where $\boldsymbol{\mu}=(\mu_1,\mu_2,\ldots,\mu_{2d})\in\{-1,1\}^{2d}$. Any arrangement of eigenvalues $\boldsymbol{\mu}$ corresponds to eigenvalue $\lambda=\sum_{i=1}^{2d}\mu_i$. Consequent, projector onto eigenspace zero of Λ reads

$$\mathbb{P}^{\Lambda}_0=\sum_{\substack{\boldsymbol{\mu}\in\{-1,1\}^{2d}\\ \sum_{i=1}^{2d}\mu_i=0}}\mathbb{P}_{\boldsymbol{\mu}}. \tag{E48}$$

Expanding each of the projectors $\mathbb{P}_{\boldsymbol{\mu}}$ into sum of products of Majorana monomials gives

$$\mathbb{P}_{\boldsymbol{\mu}}=\frac{1}{2^{2d}}\sum_{k=0}^{2d}\sum_{\mathcal{X}\in\binom{[2d]}{k}}\boldsymbol{\mu}^{\mathcal{X}}\prod_{i\in\mathcal{X}}m_i\otimes m_i, \tag{E49}$$

where we define $\boldsymbol{\mu}^{\mathcal{X}}=\prod_{i\in\mathcal{X}}\mu_i$. Inserting this expression to Eq. (E48) gives

$$\mathbb{P}^{\Lambda}_0=\frac{1}{2^{2d}}\sum_{k=0}^{2d}\sum_{\mathcal{X}\in\binom{[2d]}{k}}A_{\mathcal{X}}\prod_{i\in\mathcal{X}}m_i\otimes m_i, \tag{E50}$$

with

$$A_{\mathcal{X}}=\sum_{\substack{\boldsymbol{\mu}\in\{-1,1\}^{2d}\\ \sum_{i=1}^{2d}\mu_i=0}}\boldsymbol{\mu}^{\mathcal{X}}. \tag{E51}$$

Every $\boldsymbol{\mu}\in\{-1,1\}^{2d}$ can be identified with a subset $\mathcal{Y}_{\boldsymbol{\mu}}\subset[2d]$ defined by $\mathcal{Y}_{\boldsymbol{\mu}}=\{i\mid\mu_i=-1\}$. Under this identification $\boldsymbol{\mu}^{\mathcal{X}}=(-1)^{|\mathcal{X}\cap\mathcal{Y}_{\boldsymbol{\mu}}|}$. Consequently we obtain

$$A_{\mathcal{X}}=\sum_{\mathcal{Y}\in\binom{[2d]}{d}}(-1)^{|\mathcal{X}\cap\mathcal{Y}|}. \tag{E52}$$

Let us first observe that because $(-1)^{|\mathcal{X}\cap\mathcal{Y}|}=(-1)^d(-1)^{|\bar{\mathcal{X}}\cap\mathcal{Y}|}$, for $\bar{\mathcal{X}}=[2d]\setminus\mathcal{X}$ and $|Y|=d$, we have $C_{\mathcal{X}}=(-1)^dC_{\bar{\mathcal{X}}}$. Assuming $|X|=k\leq d$ we get

$$A_{\mathcal{X}}=\sum_{l=0}^{k}(-1)^l\sum_{\substack{\mathcal{Y}\in\binom{[2d]}{d}\\ |\mathcal{X}\cap\mathcal{Y}|=l}}=\sum_{l=0}^{k}(-l)^l\binom{k}{l}\binom{2d-k}{d-l}, \tag{E53}$$

where to get the second equality we count the number of sets $\mathcal{Y}\in\binom{[2d]}{d}$ satisfying $|\mathcal{X}\cap\mathcal{Y}|=l$, where $|X|=k\leq d$. Since we know $A_{\mathcal{X}}$ depends only on $|\mathcal{X}|$ we use the notation denoting $A_{\mathcal{X}}=A_{|\mathcal{X}|}$. Using simple algebra we obtain

$$\begin{aligned}A_k&=\sum_{l=0}^{k}(-1)^l\binom{k}{l}\binom{2d-k}{d-l}\\ &=\frac{k!(2d-k)!}{(d!)^2}\sum_{l=0}^{k}(-1)^l\binom{d}{l}\binom{d}{k-l}.\end{aligned} \tag{E54}$$

This can be further simplified using the identity

$$\sum_{l=0}^{k}(-1)^l\binom{d}{l}\binom{d}{k-l}=\begin{cases}(-1)^{k/2}\binom{d}{k/2} & \text{if } k \text{ is even}\\ 0 & \text{if } k \text{ is odd.}\end{cases} \tag{E55}$$

Denoting $A_{2p}=C_p$ and using $C_{\mathcal{X}}=(-1)^dC_{\bar{\mathcal{X}}}$ we observe that $C_{d-k}=C_k$. Inserting the expression for $A_{\mathcal{X}}$ to Eq. (E50) we finally obtain the desired result:

$$\mathbb{P}^{\Lambda}_0=\frac{1}{2^{2d}}\sum_{p=0}^{d}C_p\sum_{\mathcal{X}\in\binom{[2d]}{2p}}\prod_{i\in\mathcal{X}}m_i\otimes m_i, \tag{E56}$$

where C_p satisfy $C_p=(-1)^dC_{d-p}$ and for $p\leq\lfloor d/2\rfloor$

$$C_p=(-1)^p\frac{(2p)!(2d-2p)!}{(d!)^2}\binom{d}{p}. \tag{E57}$$

We thus establish $\mathbb{P}_{\text{act}}=\mathbb{P}_{+}\otimes\mathbb{P}_{+}\mathbb{P}_0\mathbb{P}_{+}\otimes\mathbb{P}_{+}$. The dimension of the subspace on which $\mathbb{P}_{\text{act}}$ projects, $|\tilde{\mathcal{H}}_{\text{act}}|$, can now be computed as $\text{tr}(\mathbb{P}_{\text{pas}})$ by using standard algebraic properties of Majorana operators. ■

In order to prove the following lemma we use explicit form of $\mathbb{P}_{\mathrm{pas}}$ to compute $\mathrm{tr}(\mathbb{P}_{\mathrm{pas}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}})$.

Lemma 16: *Consider the setting of our quantum advantage proposal, i.e., $d=4N$ and $n=2N$. Let $\Psi_{\mathrm{in}}\in\mathcal{D}\left[\mathcal{H}^{+}_{\mathrm{Fock}}(\mathbb{C}^{4N})\right]$. Let $\mathbb{P}_{\mathrm{act}}$ be a projector specified in Lemma 14. We then have*

$$\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}})=\frac{1}{2^{8N}}\times\left[2\sum_{q}^{N-1}C_{2q}\sum_{l=0}^{\lfloor\frac{q}{2}\rfloor}\frac{N!}{l!(q-2l)!(N-q+l)!}14^{q-2l}+C_{2N}\sum_{l=0}^{N}\frac{N!}{(l!)^2(4N-2l)!}14^{N-2l}\right]. \tag{E58}$$

where

$$C_{2q}=\frac{(4q)!(8N-4q)!}{[(4N)!]^2}\binom{4N}{2q}. \tag{E59}$$

Proof. We start be observing that due to FLO invariance of $\mathbb{P}_{\mathrm{act}}$ we have $\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}})=\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi'_{\mathrm{in}}\otimes\Psi'_{\mathrm{in}})$, where $\Psi'_{\mathrm{in}}=V\Psi_{\mathrm{in}}V^{\dagger}$, for $V\in\mathcal{G}_{\mathrm{act}}$. Note that by applying $V=\prod_{1=1}^{N}m_{3i-1}m_{4i-i}$ we can transform $|\Psi_{\mathrm{in}}\rangle=|\Psi_4\rangle^{\otimes N}$ [recall that $|\Psi_4\rangle=1/\sqrt{2}(|0011\rangle+|1100\rangle)$ into $\left|\Psi'_{\mathrm{in}}\right\rangle=|a_8\rangle^{\otimes N}$, where $|a_8\rangle=1/\sqrt{2}(|0000\rangle+|1111\rangle)$] is the state that is used, considered previously by Bravyi in the context of magic state injection for the model of computation based on Ising anyons [58] (see also Refs. [72,73]). The state $\left|a_8^{\mathcal{B}}\right\rangle\left\langle a_8^{\mathcal{B}}\right|$ on octet of normally ordered Majorana modes denoted by $\mathcal{B}\subset[2d]$ can be decomposed using Majorana monomials

$$\left|a_8^{\mathcal{B}}\right\rangle\left\langle a_8^{\mathcal{B}}\right|=\frac{1}{2^4}\left(\mathbb{I}+Q^{\mathcal{B}}+A_1^{\mathcal{B}}+A_2^{\mathcal{B}}+\ldots+A_{14}^{\mathcal{B}}\right), \tag{E60}$$

where $Q^{\mathcal{B}}=\prod_{i\in\mathcal{B}}m_i$ and operators $A_i^{\mathcal{B}}$, $i=1,\ldots,14$ are quartic (i.e., fourth other) Majorana monomials supported on modes belonging to $\mathcal{B}$ and satisfying $(A_i^{\mathcal{B}})^2=\mathbb{I}$. We do not need the explicit form of $|a_8\rangle\langle a_8|$ but it can be found in the works cited above. The algebraic framework of Majorana fermion operators allows us to write the equivalent input state $\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|$ as a product (in a standard operator sense) of states $|a_8\rangle\langle a_8|$ supported on disjoint octets of modes

$$\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|=\prod_{i=1}^{N}\left|a_8^{\mathcal{B}_i}\right\rangle\left\langle a_8^{\mathcal{B}_i}\right|, \tag{E61}$$

where $\mathcal{B}_1=\{1,2,\ldots,8\}$, $\mathcal{B}_1=\{1,2,\ldots,8\}$, $\mathcal{B}_2=\{9,10,\ldots,16\}$, etc. We proceed similarly as in the proof of Proposition 4 and expand the above expressions into product of Majorana monomials and obtain

$$\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|=\frac{1}{2^{4N}}\sum_{\mathcal{X}\in\mathcal{C}_{A8}}(-1)^{F(\mathcal{X})}\prod_{i\in\mathcal{X}}m_i, \tag{E62}$$

where $\mathcal{C}_{A8}$ is a collection of subsets of $8N$ Majorana modes that appear in the product expansion of $\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|$ and $(-1)^{F(\mathcal{X})}$ a sign possibly depending on a subset $\mathcal{X}$. Because $\left|\Psi'_{\mathrm{in}}\right\rangle\in\mathcal{H}^{+}_{\mathrm{Fock}}(\mathbb{C}^d)$ and the form projector $\mathbb{P}_{\mathrm{act}}$ [cf. Eq. (E41)] we have $\mathrm{tr}\left(\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|^{\otimes 2}\mathbb{P}_{\mathrm{act}}\right)=\mathrm{tr}\left(\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|^{\otimes 2}\mathbb{P}_0\right)$, where $\mathbb{P}_0$ is given in Eq. (E42). Combining Eq. (E62) with Eq. (E42) gives

$$\mathrm{tr}\left(\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|^{\otimes 2}\mathbb{P}_0\right)=\frac{1}{2^{16N}}\sum_{p=0}^{4N}C_p\sum_{\mathcal{Z}\in\binom{[8N]}{2p}}\sum_{\mathcal{X},\mathcal{Y}\in\mathcal{C}_{A8}}(-1)^{F(\mathcal{X})+F(\mathcal{Y})}\,\mathrm{tr}\left[\prod_{i\in\mathcal{X}}m_j\otimes\prod_{k\in\mathcal{Y}}m_i\prod_{k\in\mathcal{Z}}m_k\otimes m_k\right]. \tag{E63}$$

Using

$$(-1)^{F(\mathcal{X})+F(\mathcal{Y})}\,\mathrm{tr}\left[\prod_{i\in\mathcal{X}}m_j\otimes\prod_{k\in\mathcal{Y}}m_i\prod_{k\in\mathcal{Z}}m_k\otimes m_k\right]=2^{8N}\delta_{\mathcal{X},\mathcal{Y}}\delta_{\mathcal{Z},\mathcal{X}} \tag{E64}$$

we obtain

$$\mathrm{tr}\left(\left|\Psi'_{\mathrm{in}}\right\rangle\left\langle\Psi'_{\mathrm{in}}\right|^{\otimes 2}\mathbb{P}_0\right)=\frac{1}{2^{8N}}\sum_{p=0}^{4N}C_p\sum_{\mathcal{X}\in\binom{[8N]}{2p}\cap\mathcal{C}_{A8}}. \tag{E65}$$

Recall that from the definition of $\mathcal{C}_{A8}$, this collection of subsets of $[8N]$ consists only of subsets that have cardinality divisible by 4. Therefore, the above can be described equivalently by

$$\operatorname{tr}\left(\left|\Psi'_{\text{in}}\right\rangle\left\langle\Psi'_{\text{in}}\right|^{\otimes 2}\mathbb{P}_0\right)=\frac{1}{2^{8N}}\sum_{q=0}^{2N}C_{2q}\sum_{\mathcal{X}\in\binom{[8N]}{4q}\cap\mathcal{C}_{A8}}. \tag{E66}$$

Moreover, from $Q\left|\Psi'_{\text{in}}\right\rangle\left\langle\Psi'_{\text{in}}\right|=\left|\Psi'_{\text{in}}\right\rangle\left\langle\Psi'_{\text{in}}\right|$ we get $\bar{\mathcal{X}}\in\mathcal{C}_{A8}$ if and only if $\mathcal{X}\in\mathcal{C}_{A8}$ and consequently

$$\sum_{\mathcal{X}\in\binom{[8N]}{4q}\cap\mathcal{C}_{A8}}=\sum_{\mathcal{X}\in\binom{[8N]}{8N-4q}\cap\mathcal{C}_{A8}}. \tag{E67}$$

Using this and the property $C_{2q}=C_{8N-2q}$ we finally get

$$\operatorname{tr}\left(\left|\Psi'_{\text{in}}\right\rangle\left\langle\Psi'_{\text{in}}\right|^{\otimes 2}\mathbb{P}_0\right)=\frac{1}{2^{8N}}\sum_{q=0}^{N-1}2\left(C_{2q}\sum_{\mathcal{X}\in\binom{[8N]}{4q}\cap\mathcal{C}_{A8}}\right)+\left(C_{2N}\frac{1}{2^{8N}}\sum_{\mathcal{X}\in\binom{[8N]}{4N}\cap\mathcal{C}_{A8}}\right). \tag{E68}$$

Therefore, we reduce the problem of computing $\operatorname{tr}\left(\left|\Psi'_{\text{in}}\right\rangle\left\langle\Psi'_{\text{in}}\right|^{\otimes 2}\mathbb{P}_0\right)$ [equal to $\operatorname{tr}(|\Psi_{\text{in}}\rangle\langle\Psi_{\text{in}}|^{\otimes 2}\mathbb{P}_0)$] to the problem of counting different sets of cardinality $4q$ ($q=0,1,\ldots,N$) one can find in $\mathcal{C}_{A8}$. This problem can be tackled using similar technique to the one used in the proof of Proposition 4, i.e., by introducing the *pattern of types* of subsets in $\mathcal{C}_{A8}$. We have $8N$ Majorana modes in total. In what follows we refer to "standard octets" as N disjoint octets on which states $\left|a_8^{\mathcal{B}_i}\right\rangle$ are supported in Eq. (E61). A subset $\mathcal{X}\in\mathcal{C}_{A8}$ satisfying $|\mathcal{X}|=4q$ ($q\leq N$) can be characterized, in analogy to Eq. (E32), by the pattern of types, i.e., a triple

$$\mathcal{X}\longmapsto\mathcal{L}(\mathcal{X})=\left(l_{\text{oct}}[\mathcal{X}],\ l_{\text{empty}}[\mathcal{X}],\ l_{\text{quad}}[\mathcal{X}]\right), \tag{E69}$$

where $l_{\text{oct}}[\mathcal{X}]$ counts the number of standard octets contained in $\mathcal{X}$, $l_{\text{empty}}[\mathcal{X}]$ counts how many standard octets are not populated by elements of $\mathcal{X}$, and finally $l_{\text{quad}}[\mathcal{X}]$ is the number of octets in which $\mathcal{X}$ intersects only in four elements [note that from the construction of $\mathcal{C}_{A8}$ and due to specific form of the state $\left|a_8^{\mathcal{B}}\right\rangle\left\langle a_8^{\mathcal{B}}\right|$ in Eq. (E60) these are the only possibilities]. With these concepts counting of sets $\mathcal{X}\in\mathcal{C}_{A8}$ of carnality $4q$ can be done analogously as in Proposition 4, i.e., by counting how many sets of different "pattern of types" $\mathcal{L}(\mathcal{X})$ of given carnality there are. The final results reads

$$\left|\left\{\mathcal{X}\mid\mathcal{X}\in\binom{[8N]}{4q},\ \mathcal{X}\in\mathcal{C}_{A8}\right\}\right|=\sum_{l=0}^{\lfloor\frac{q}{2}\rfloor}\frac{N!}{l!(q-2l)!(N-q+l)!}14^{q-2l}, \tag{E70}$$

where l labels the number of possible "fully occupied" standard octets in $\mathcal{X}$ of carnality $4q$. The term 14^{q-2l} appears because for the said value of fully occupied octets there are necessarily $q-2l$ octets of quartic type, and every such octet there is exactly 14 possibilities. We conclude the proof by using the above identity in Eq. (E68) and employing the explicit formula for C_{2q} from Eq. (E43). ■

3. Computation of the sums

In this part we prove the bounds on the combinatorial sums appearing in Lemma 13, Lemma 15. This ultimately proves anticoncentration bounds for passive and active FLO circuits initialized in magic input states Ψ_{in} in Theorem 1.

Our general strategy for the analytical part is based on the following tight inequalities satisfied by binomial and trinomial coefficients.

Lemma 17: (Bounds for binomial and trinomial coefficients.) *Let n,k be a natural numbers such that $k\in\{1,\ldots,n-1\}$. Let $x=k/n$. Then we have*

$$c\cdot\sqrt{\frac{n}{k(n-k)}}\exp\left[n\,h(x)\right]\leq\binom{n}{k}\leq C\cdot\sqrt{\frac{n}{k(n-k)}}\exp\left[n\,h(x)\right], \tag{E71}$$

where $c=1/2\sqrt{2}$, $C=1/\sqrt{2\pi}$, and $h(x)=-x\log(x)-(1-x)\log(1-x)$ is the binary entropy.

Moreover, let k,l,m be nonzero natural numbers such that $k+l+m=n$. Let $x=k/n, y=l/n, z=m/n$. Then we have

$$a\sqrt{\frac{n}{k\cdot l\cdot m}}\exp\left[n\,h(x,y,z)\right]\leq\binom{n}{k,l,m}\leq A\sqrt{\frac{n}{k\cdot l\cdot m}}\exp\left[n\,h(x,y,z)\right], \tag{E72}$$

where $a=1/8$, $A=1/2\pi$ and $h(x,y,z)=-x\log(x)-y\log(y)-z\log(z)$ is the entropy of three-outcome probability distribution.

The inequality (E71) can be found in Lemma 7 in Chapter 10 of Ref. [118] while Eq. (E72) follows from it due to identity $\binom{n}{k,l,m}=\binom{n}{k}\binom{l+m}{m}$.

We first consider the case of passive FLO. We observe that from Eq. (E38) it follows that

$$\mathrm{tr}\left(\mathbb{P}_{\mathrm{pas}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}}\right)\leq\frac{2}{2N+1}\sum_{k=0}^{N}\sum_{l=0}^{\lfloor k/2\rfloor}\frac{\binom{N}{l,k-2l,N-k+l}}{\binom{2N}{k}}. \tag{E73}$$

Lemma 18: *Consider the setting of our quantum advantage proposal, i.e., $d=4N$ and $n=2N$. Let $\Psi_{\mathrm{in}}\in\mathcal{D}[\bigwedge^{2N}(\mathbb{C}^{4N})]$. Let $\mathbb{P}_{\mathrm{pas}}$ be defined as in Lemma 12. We then have*

$$\mathrm{tr}\left(\mathbb{P}_{\mathrm{pas}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}}\right)\leq\frac{C_{\mathrm{pas}}}{N},\ \textit{for}\ C_{\mathrm{pas}}=5.7. \tag{E74}$$

Proof. Let us denote

$$f_N(k,l):=\frac{\binom{N}{l,k-2l,N-k+l}}{\binom{2N}{k}}. \tag{E75}$$

From Eq. (E73) it follows that

$$\mathrm{tr}\left(\mathbb{P}_{\mathrm{pas}}\Psi_{\mathrm{in}}\otimes\Psi_{\mathrm{in}}\right)\leq\frac{2}{2N+1}\sum_{k=0}^{N}\sum_{l=0}^{\lfloor k/2\rfloor}f_N(k,l)\leq\frac{1}{N}\times\left(\mathcal{A}_{k=0}+\mathcal{A}_{l=0}+\mathcal{A}_{k=2l}+\mathcal{A}_{\mathrm{gen}}\right), \tag{E76}$$

where

$$\mathcal{A}_{k=0}=f_N(0,0)=1, \tag{E77}$$

$$\mathcal{A}_{l=0}=\sum_{k=1}^{N}\frac{\binom{N}{k}}{\binom{2N}{k}}, \tag{E78}$$

$$\mathcal{A}_{k=2l}=\sum_{\substack{k>1\\k\ \mathrm{even}}}^{N}\frac{\binom{N}{k/2}}{\binom{2N}{k}}, \tag{E79}$$

$$\mathcal{A}_{\mathrm{gen}}=\sum_{k=1}^{N}\sum_{l=1}^{l<k/2}f_N(k,l). \tag{E80}$$

We upper bound each term above separately (except for the trivial case of $\mathcal{A}_{k=0}$). The following analytical proof for the bound requires $N\geq 130$. In particular, the bound for Eq. (E85) $\mathcal{A}_{l=0}$ is valid for $N\geq 40$, and the bound (E101) for $\mathcal{A}_{\mathrm{gen}}$ is valid for $N\geq 130$. At the end of the proof, we show in Fig. 6 that the bound also holds for all smaller values of N.

Upper bound on $\mathcal{A}_{l=0}$. In this case, we derive a bound valid for $N>40$. The bounds from Lemma 16 give

$$\mathcal{A}_{l=0}\leq\frac{1}{\binom{2N}{N}}+\frac{C}{c}\sum_{k=1}^{N-1}\sqrt{\frac{2N-k}{2(N-k)}}\exp\times\left[N\left\{h(k/N)-2h(k/2N)\right\}\right]. \tag{E81}$$

We use now the inequality $h(x)-2h(x/2)\leq -2/3x$, valid for $x\in[0,1]$ to obtain

$$\mathcal{A}_{l=0}\leq\frac{1}{\binom{2N}{N}}+\frac{C}{c}\sum_{k=1}^{N-1}\sqrt{\frac{2N-k}{2(N-k)}}\exp\left(-\frac{2k}{3}\right). \tag{E82}$$

We then apply the bound $\binom{2N}{N}\geq c2^{2N}\sqrt{2/N}$ and divide the sum over k into two parts

$$\mathcal{A}_{l=0}\leq\frac{\sqrt{N}}{\sqrt{2}c}2^{-2N}+\frac{C}{c}\left[\sum_{k=1}^{k\leq 1/2N}\sqrt{\frac{2N-k}{2(N-k)}}\exp\left(-\frac{2k}{3}\right)+\sum_{k>1/2N}^{N-1}\sqrt{\frac{2N-k}{2(N-k)}}\exp\left(-\frac{2k}{3}\right)\right]. \tag{E83}$$

For $k\leq N/2$ we have $\sqrt{2N-k/2(N-k)}\leq\sqrt{3/2}$ and therefore

$$\mathcal{A}_{l=0}\leq\frac{\sqrt{N}}{\sqrt{2}c}2^{-2N}+\frac{C}{c}\left[\sqrt{\frac{3}{2}}\frac{1}{e^{2/3}-1}+\frac{N^{3/2}}{2}\exp\left(-\frac{N}{3}\right)\right], \tag{E84}$$

where we utilize the expression for the sum of geometric progression and the upper bound $\sqrt{2N-k/2(N-k)}\leq\sqrt{N}$, valid for $k\leq N-1$. Using expression (E84) it is easy to verify that for $N>40$ we have

$$\mathcal{A}_{l=0}\leq\frac{3}{2}. \tag{E85}$$

Upper bound on $\mathcal{A}_{k=2l}$. Estimates for binomials from Lemma 16 yield

$$\mathcal{A}_{k=2l}\leq\frac{C\sqrt{2}}{c}\sum_{\substack{k>1\\k\ \mathrm{even}}}^{N}\exp\left[-Nh(k/2N)\right]. \tag{E86}$$

Concavity of binary entropy $h(\cdot)$ implies that for $x \in [0,1]$ we have $[\log(2)/2]x \leq h(x/2)$ and consequently

$$\mathcal{A}_{k=2l} \leq \frac{C\sqrt{2}}{c} \sum_{\substack{k>1 \\ k \text{ even}}}^{N} \exp\left(-\frac{k\log(2)}{2}\right) = \frac{C\sqrt{2}}{c} \sum_{p=1}^{\lfloor N/2 \rfloor} 2^{-p}. \tag{E87}$$

The sum of the geometric series in the above expression is upper bounded by 1 and therefore

$$\mathcal{A}_{k=2l} \leq \frac{C\sqrt{2}}{c} \leq \frac{8}{5}. \tag{E88}$$

Upper bound on $\mathcal{A}_{\text{gen}}$. In the following proof, we require that $N \geq 130$. For the generic points in the sum (E73) inequalities from Lemma 16 give

$$\mathcal{A}_{\text{gen}} \leq \frac{A}{\sqrt{2}c} \sum_{k=1}^{N} \sum_{l=1}^{l<k/2} \sqrt{\frac{k(2N-k)}{l(k-2l)(N-k+l)}} \exp\left(N\left\{h\left[x_l, y_k - 2x_l, 1 - y_k + x_l\right] - 2h\left[y_k/2\right]\right\}\right), \tag{E89}$$

where $x_l = l/N$, $y_k = k/N$. Note that $k=1$ and $k=2$ are implicitly excluded from the above sum because of the constraints on l and hence

$$\mathcal{A}_{\text{gen}} \leq \frac{A}{\sqrt{2}c} \sum_{k=3}^{N} \sum_{l=1}^{l<k/2} \sqrt{\frac{k(2N-k)}{l(k-2l)(N-k+l)}} \exp\left(N\left\{h\left[x_l, y_k - 2x_l, 1 - y_k + x_l\right] - 2h\left[y_k/2\right]\right\}\right). \tag{E90}$$

In order to upper bound the expression we maximize the function

$$F(x,y) = h\left(x, y-2x, 1-y+x\right) - 2h\left(y/2\right) \tag{E91}$$

over $x \in [0, y/2]$, for fixed value of $y \in [0,1]$. Looking for critical points reduces the problem to solving quadratic equation, which gives a unique solution in the interval $[0, y/2]$:

$$x_{\text{opt}}(y) = \frac{1}{6}\left(1 + 3y - \sqrt{1 + 6y - 3y^2}\right). \tag{E92}$$

Crucially, the function $F_{\text{opt}}(y) := F(x_{\text{opt}}(y), y)$ is a continuous function of parameter y, which is also analytic in the interior the interval $(0,1)$. Moreover, $F_{\text{opt}}(y)$ satisfies (see Fig. 11):

$$F_{\text{opt}}(y) \leq -\frac{1}{2}y \text{ for } y \in [0, 1/3],$$
$$F_{\text{opt}}(y) \leq -\frac{1}{4}y \text{ for } y \in [0, 1]. \tag{E93}$$

It follows that

$$N\left(h\left[x_l, y_k - 2x_l, 1 - y_k + x_l\right] - 2h\left[y_k/2\right]\right) \leq -\frac{1}{2}k \text{ for } 1 \leq k \leq N/3, \tag{E94}$$

$$N\left(h\left[x_l, y_k - 2x_l, 1 - y_k + x_l\right] - 2h\left[y_k/2\right]\right) \leq -\frac{1}{4}k \text{ for } 1 \leq k \leq N. \tag{E95}$$

Moreover, for integer l satisfying $1 \leq l < k/2$ we have $l(k-2l) \geq (k-2)/2$ and consequently for $k \geq 3$ we have $k/l(k-2l) \leq 2k/k - 2 \leq 6$. As a result we have

$$\sum_{l=1}^{l<k/2} \sqrt{\frac{k(2N-k)}{l(k-2l)(N-k+l)}} \leq \frac{\sqrt{6}k}{2} \sqrt{\frac{2N-k}{N-k+1}}. \tag{E96}$$

Inserting Eqs. (E94) and (E96) into Eq. (E90) gives

$$\mathcal{A}_{\text{gen}} \leq \frac{\sqrt{3}A}{2c}\left[\sum_{k=3}^{k\leq N/3} \sqrt{\frac{2N-k}{N-k+1}}\, k \exp\left(-\frac{k}{2}\right) + \sum_{k>N/3}^{N} \sqrt{\frac{2N-k}{N-k+1}}\, k \exp\left(-\frac{k}{4}\right)\right]. \tag{E97}$$

Observing that for $k \leq N/3$ we have $\sqrt{2N-k/N-k+1} \leq \sqrt{5/2}$, while and for general $k \leq N$ $\sqrt{2N-k/N-k+1} \leq \sqrt{N}$, we obtain

$$\mathcal{A}_{\text{gen}} \leq \frac{\sqrt{3}A}{2c}\left[\sqrt{\frac{5}{2}} \sum_{k=3}^{k\leq N/3} k \exp\left(-\frac{k}{2}\right) + \frac{2N^{\frac{3}{2}}}{3} \exp\left(-\frac{N}{12}\right)\right]. \tag{E98}$$

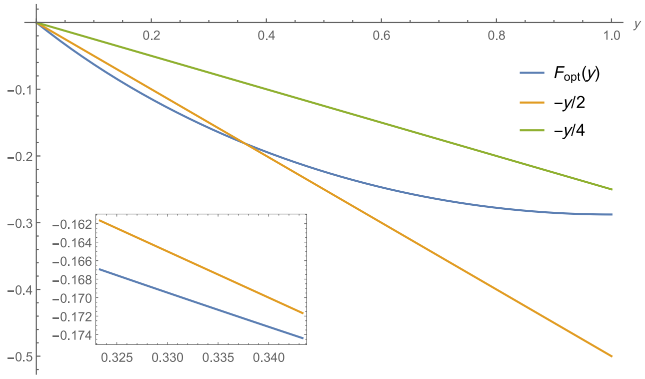


FIG. 11. Function $F_{\mathrm{opt}}(y) := F(x_{\mathrm{opt}}(y), y)$ where F is defined in Eq. (E91) and $x_{\mathrm{opt}}(y)$ is given in Eq. (E92). The function is bounded by $-y/3$ in the interval $[0, 1/3]$ and by $-y/4$ in the interval $[1/3, 1]$. The inset plot shows that the inequality is also valid near $y = 1/3$.

We bound the first summand as follows:

$$\sum_{k=3}^{k\leq N/3} k \exp\left(-\frac{k}{2}\right) \leq \sum_{k=3}^{\infty} k \exp\left(-\frac{k}{2}\right) = \frac{3\sqrt{e}-2}{(\sqrt{e}-1)^2 e}. \tag{E99}$$

This finally gives us

$$\mathcal{A}_{\mathrm{gen}} \leq \frac{\sqrt{15}A}{2\sqrt{2}c}\frac{3\sqrt{e}-2}{(\sqrt{e}-1)^2 e} + \frac{A}{\sqrt{3}c} N^{\frac{3}{2}} \exp\left(-\frac{N}{12}\right). \tag{E100}$$

Using the above expression we get that for $N \geq 130$ we have

$$\mathcal{A}_{\mathrm{gen}} \leq \frac{8}{5}. \tag{E101}$$

Finally, combining bounds (E85), (E88), and (E101) together with $\mathcal{A}_{k=0} = 1$ we see that for $N \geq 130$,

$$\mathcal{A}_{k=0} + \mathcal{A}_{l=0} + \mathcal{A}_{k=2l} + \mathcal{A}_{\mathrm{gen}} \leq 5.7. \tag{E102}$$

Inserting this into the bound (E76) proves the lemma for $N \geq 130$. For $N \leq 130$, the validity of the bound can be verified numerically as shown in Fig. 6, which completes the proof. ■

Analogously for the active FLO case, Eq. (E58) implies that

$$\mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}} \otimes \Psi_{\mathrm{in}}) \leq \frac{\binom{8N}{4N}}{2^{8N-1}} \sum_{q=0}^{N} \sum_{l=0}^{\lfloor \frac{q}{2} \rfloor} \frac{\binom{4N}{2q}\binom{N}{l,q-2l,N-q+l}}{\binom{8N}{4q}} 14^{q-2l}. \tag{E103}$$

Lemma 19: *Consider the setting of our quantum advantage proposal, i.e., $d = 4N$ and $n = 2N$. Let $\Psi_{\mathrm{in}} \in \mathcal{D}[\bigwedge^{2N}(\mathbb{C}^{4N})]$. Let $\mathbb{P}_{\mathrm{act}}$ be defined as in Lemma 14. We then have*

$$\mathrm{tr}\,(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}} \otimes \Psi_{\mathrm{in}}) \leq \frac{C_{\mathrm{act}}}{\sqrt{\pi N}}, \textit{ for } C_{\mathrm{act}} = 16.2. \tag{E104}$$

Proof. Our proof strategy is analogous to the one used in the case of passive FLO. Let us denote

$$g_N(q,l) := \frac{\binom{4N}{2q}\binom{N}{l,q-2l,N-q+l}}{\binom{8N}{4q}} 14^{q-2l}. \tag{E105}$$

It follows from Eq. (E103) and the entropic bound for binomial coefficients in Lemma 16,

$$\frac{\binom{8N}{4N}}{2^{8N-1}} \leq \frac{1}{\sqrt{\pi N}}, \tag{E106}$$

that

$$\begin{aligned} \mathrm{tr}(\mathbb{P}_{\mathrm{act}}\Psi_{\mathrm{in}} \otimes \Psi_{\mathrm{in}}) &\leq \frac{1}{\sqrt{\pi N}} \sum_{q=0}^{N} \sum_{l=0}^{\lfloor \frac{q}{2} \rfloor} g_N(q,l) \\ &\leq \frac{1}{\sqrt{\pi N}} (\mathcal{B}_{q=0} + \mathcal{B}_{l=0} + \mathcal{B}_{q=2l} + \mathcal{B}_{\mathrm{gen}}), \end{aligned} \tag{E107}$$

where

$$\mathcal{B}_{q=0} = g_N(0,0) = 1, \tag{E108}$$

$$\mathcal{B}_{l=0} = \sum_{q=1}^{N} \frac{\binom{4N}{2q}\binom{N}{q}}{\binom{8N}{4q}} 14^{q}, \tag{E109}$$

$$\mathcal{B}_{q=2l} = \sum_{\substack{q>1 \\ q \text{ even}}}^{N} \frac{\binom{4N}{2q}\binom{N}{q/2}}{\binom{8N}{4q}}, \tag{E110}$$

$$\mathcal{B}_{\mathrm{gen}} = \sum_{q=1}^{N} \sum_{l=1}^{l<q/2} g_N(q,l). \tag{E111}$$

We upper bound each term above separately (except for the trivial case of $\mathcal{B}_{q=0}$). The following analytical proof for the bound requires $N \geq 7000$. In particular, the bound for Eq. (E85) $\mathcal{B}_{l=0}$ is valid for $N \geq 1000$, and the bound (E101) for $\mathcal{B}_{\mathrm{gen}}$ is valid for $N \geq 7000$. At the end of the proof, we show in Fig. 7 that the bound (E104) also holds for all smaller values of $N \leq 7000$ by numerically evaluating right-hand side of Eq. (E103).

Upper bound on $\mathcal{B}_{l=0}$. For this term, we require that $N \geq 1000$. The entropic bound in Lemma 16 implies that

$$\mathcal{B}_{l=0} \leq \frac{\binom{4N}{2N}}{\binom{8N}{4N}}14^N + \frac{C^2\sqrt{2}}{c}\sum_{q=1}^{N-1}\sqrt{\frac{N}{q(N-q)}}\exp \times [N\{h(q/N) - 4h(q/2N) + \log(14)q/N\}]. \tag{E112}$$

To upper bound the sum, we split the sum into two sums: one from $q = 1$ to $q \leq N/5$ and another from $q > N/5$ to $q = N - 1$, and upper bound the function

$$H(x) := h(x) - 4h(x/2) + x\log(14), \tag{E113}$$

$x \in [0, 1]$ in the intervals $[0, 1/5]$ and $(1/5, 1]$ separately. In particular, we have that (see also Fig. 12)

$$H(x) \leq -\frac{4}{3}x \text{ for } x \in [0, 2/5],$$
$$H(x) \leq -\frac{1}{18}x \text{ for } x \in [0, 1]. \tag{E114}$$

Together with the bound $\binom{4N}{2N}/\binom{8N}{4N} \leq C\sqrt{2}/c2^{-4N}$ and $\sqrt{N/[q(N-q)]} \leq \sqrt{2}$ valid for $N \geq 2$ (this is because $\sqrt{N/q(N-q)}$ is convex for $q \in [1, N-1]$ and thus the expression takes the maximum values at the end points), we obtain

$$\mathcal{B}_{l=0} \leq \frac{C\sqrt{2}}{c}\left(\frac{14}{16}\right)^N + \frac{2C^2}{c} \times \left(\sum_{q=1}^{q\leq N/5}\exp(-4q/3) + \sum_{q>N/5}^{N-1}\exp(-q/18)\right) \tag{E115}$$

$$\leq \frac{C\sqrt{2}}{c}\left(\frac{14}{16}\right)^N + \frac{2C^2}{c}\left(\frac{1}{e^{4/3}-1} + \frac{4N}{5}\exp \times \left[-\frac{N}{18\cdot 5}\right]\right), \tag{E116}$$

where we use the sum of the geometric series to arrive at the final expression. Using the expression (E116), it can be verified that

$$\mathcal{B}_{l=0} \leq \frac{1}{3} \tag{E117}$$

holds for $N \geq 1000$.

Upper bound on $\mathcal{B}_{q=2l}$. From Lemma 16 we see that

$$\mathcal{B}_{q=2l} \leq \frac{C^2\sqrt{2}}{c}\sum_{\substack{q>1\\ q \text{ even}}}^{N}\sqrt{\frac{N}{\frac{q}{2}(N-\frac{q}{2})}}\exp[-3Nh(q/2N)]. \tag{E118}$$

Now by concavity of $h(x)$ for $x \in [0, \frac{1}{2}]$ we have $\log(2)x/2 \leq h(x/2)$ for $x \in [0, 1]$. Then

$$\mathcal{B}_{q=2l} \leq \frac{C^2\sqrt{2}}{c}\sum_{\substack{q>1\\ q \text{ even}}}^{N}\sqrt{\frac{N}{\frac{q}{2}(N-\frac{q}{2})}}\exp[-3q\log(2)/2] \tag{E119}$$

$$= \frac{C^2\sqrt{2}}{c}\sum_{p=1}^{\lfloor N/2\rfloor}\sqrt{\frac{N}{p(N-p)}}2^{-3p}. \tag{E120}$$

We can bound $\sqrt{N/p(N-p)} \leq \sqrt{2}$ the same way as in the passive case. Then we obtain

$$\mathcal{B}_{q=2l} \leq \frac{2C^2}{7c} \leq 0.13, \tag{E121}$$

where we use that the geometric sum of 2^{-3p} is bounded by $1/7$.

Upper bound on $\mathcal{B}_{\text{gen}}$. Following bounds from Lemma 16 and defining $x_l = l/N$ and $y_q = q/N$ we obtain

$$\mathcal{B}_{\text{gen}} \leq \frac{\sqrt{2}CA}{c}\sum_{q=1}^{N}\sum_{l=1}^{l<q/2}\sqrt{\frac{N}{l(q-2l)(N-q+l)}}\exp \times [NG(x_l, y_q)], \tag{E122}$$

where, following the analogous construction in Lemma 17, we introduce

$$G(x,y) := -4h(y/2) + h(x, y-2x, 1-y+x) + (y-2x)\log(14). \tag{E123}$$

As in the case of passive FLO, our strategy is to upper bound $G(x,y)$ by a function that allows for analytical treatment. To this end, we first optimize $G(x,y)$ over $x \in [0, y/2]$ for fixed $y \in [0, 1]$. Solving for the critical points gives the following optimal solution $x_{\text{opt}} \in [0, y/2]$ [at the extremal points of this interval function $G(x,y)$, treated as a function of x for fixed y, takes smaller values]

$$x_{\text{opt}}(y) = \frac{1}{96}\left(-49 + 48y + 7\sqrt{40 - 96y + 48y^2}\right). \tag{E124}$$

The maximum of $G(x,y)$ over $x \in [0, y/2]$, $G_{\text{opt}}(y) := G(x_{\text{opt}}(y), y)$ is a continuous function of $y \in [0, 1]$ and

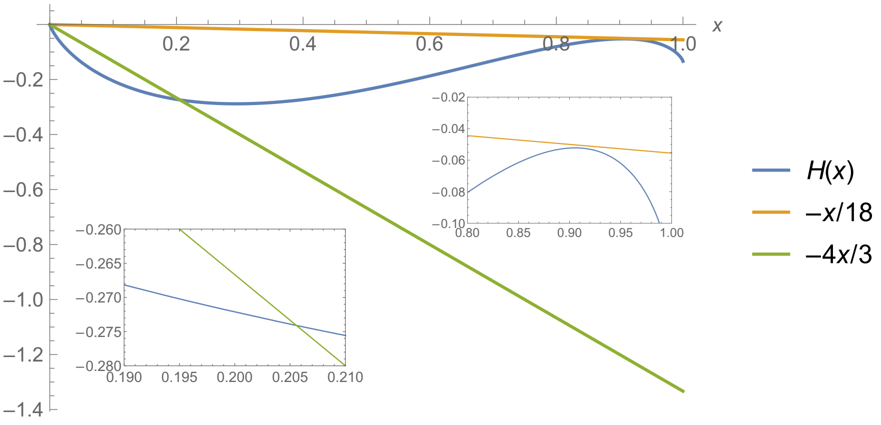


FIG. 12. Function $H(x)$ defined in Eq. (E114). The function is bounded above by $-4x/3$ in the interval $[0,1/5]$ and by $-x/18$ in the interval $[0,1]$. The inset plot shows the validity of the upper bound in each interval.

also analytic for $y\in(0,1)$. We can bound $G_{\mathrm{opt}}(y)$ in the following way (see Fig. 13):

$$\begin{aligned} G_{\mathrm{opt}}(y) &\leq -y/3 \text{ for } y\in[0,1/2]\,,\\ G_{\mathrm{opt}}(y) &\leq -y/100 \text{ for } y\in[1/5,0.925]\,,\\ G_{\mathrm{opt}}(y) &\leq -(1-y)^2 \text{ for } y\in[0.925,1]\,. \end{aligned} \tag{E125}$$

We need much more refined information about $G(x,y)$ than in the case of analogous considerations for passive FLO. Namely, we need to control how fast $G(x,y)$ decays as a function of $x-x_{\mathrm{opt}}(y)$, for fixed y. To this end we compute for $x\in(0,y/2)$, $y\in(0,1)$

$$\partial_x^2 G(x,y) = -\left(\frac{1}{x}+\frac{1}{1-y+x}+\frac{4}{y-2x}\right). \tag{E126}$$

From the above expression we get [119]

$$\partial_x^2 G(x,y)\leq -16 \text{ for } x\in(0,y/2) \text{ and } \partial_x^2 G(x,y)\leq -\frac{2}{3x_{\mathrm{opt}}(y)} \text{ for } x\in\left[\frac{x_{\mathrm{opt}}(y)}{2},\frac{3x_{\mathrm{opt}}(y)}{2}\right]. \tag{E127}$$

Using the analyticity of $G(x,y)$ as a function of x inside the interval $(0,y/2)$, we can Taylor expand it around $x_{\mathrm{opt}}(y)$ (for fixed value of y):

$$G(x,y) = G_{\mathrm{opt}}(y) + [\partial_x G(x_{\mathrm{opt}}(y),y)][x-x_{\mathrm{opt}}(y)] + \int_{x_{\mathrm{opt}}(y)}^{x} d\tau\, \partial_\tau G(\tau,y). \tag{E128}$$

Using the fact that $x_{\mathrm{opt}}(y)$ is a critical point and bounds, identity

$$\partial_\tau G(\tau,y) = \int_{x_{\mathrm{opt}}(y)}^{\tau} dx \partial_x^2 G(x,y) \tag{E129}$$

and bounds from Eq. (E127) we get finally get

$$G(x,y)\leq G_{\mathrm{opt}}(y) - 8[x-x_{\mathrm{opt}}(y)]^2 \qquad \text{for } x\in[0,y/2],\ y\in[0,1], \tag{E130}$$

$$G(x,y)\leq G_{\mathrm{opt}}(y) - \frac{1}{3x_{\mathrm{opt}}(y)}[x-x_{\mathrm{opt}}(y)]^2 \qquad \text{for } x\in\left[\frac{x_{\mathrm{opt}}(y)}{2},\frac{3x_{\mathrm{opt}}(y)}{2}\right],\ y\in[0,1]\,. \tag{E131}$$

Coming back to the bound on $\mathcal{B}_{\mathrm{gen}}$ from Eq. (E122), similarly to the case of passive FLO, due to constrains on l, the sum appearing in Eq. (E122) effectively starts from $q=3$. Moreover, we also note that $l(q-2l)\geq(q-2)/2$ and therefore

$$\sqrt{\frac{N}{l(q-2l)(N-q+l)}}\leq\sqrt{\frac{2N}{(q-2)(N-q+l)}}\leq\sqrt{\frac{2N}{N-2}}, \tag{E132}$$

where in the second inequality we use the fact that $q \in [3, N]$ and $l \geq 1$. Using the above and expanding the expression in Eq. (E122) in the different intervals defined in Eq. (E125) we obtain

$$\mathcal{B}_{\mathrm{gen}} \leq \frac{2CA}{c}\sqrt{\frac{N}{N-2}}\left(\sum_{q=3}^{q\leq N/2}\sum_{l=1}^{l<q/2}\exp[-q/2] + \sum_{q>N/2}^{q<0.925N}\sum_{l=1}^{l<q/2}\exp[-q/100]\right) \tag{E133}$$

$$+\frac{\sqrt{2}CA}{c}\sum_{q>0.925N}^{N}\sum_{l=1}^{l<q/2}\sqrt{\frac{N}{l(q-2l)(N-q+l)}}\exp[NG(x_l,y_q)]\,. \tag{E134}$$

Two sums from Eq. (E133) can be handled analogously as in the case of passive FLO:

$$\frac{2CA}{c}\sqrt{\frac{N}{N-2}}\left(\sum_{q=3}^{q\leq N/5}\sum_{l=1}^{l<q/2}\exp[-q/3] + \sum_{q>N/2}^{q<0.925N}\sum_{l=1}^{l<q/2}\exp[-q/100]\right) \tag{E135}$$

$$\leq \frac{2CA}{c}\sqrt{\frac{N}{N-2}}\left(\sum_{q=3}^{\infty}(q/2)\exp(-q/3) + +(N^{3/2}/2)\exp\left[-\frac{N}{200}\right]\right). \tag{E136}$$

$$= \frac{2CA}{c}\sqrt{\frac{N}{N-2}}\left(\frac{3e^{1/3}-2}{2e^{2/3}(e^{1/3}-1)} + (N^{3/2}/4)\exp\left[-\frac{N}{200}\right]\right) \leq 2\,, \tag{E137}$$

where the last inequality is valid for $N \geq 1800$. The sum in Eq. (E134) is analyzed using inequalities (E130) and (E131). For fixed y_q (which corresponds to $q = y_q N$) we set $l_{\mathrm{opt}}(y_q) = x_{\mathrm{opt}}(y_q)N$ and divide the range of summation over l in Eq. (E134) into two parts that corresponds to intervals in bounds (E130) and (E131), respectively,

$$\mathcal{L}_q^{\max} = \left\{ l \;\middle|\; \frac{1}{2}l_{\mathrm{opt}}(y_q) \leq l \leq \frac{3}{2}l_{\mathrm{opt}}(y_q) \right\}, \tag{E138}$$

$$\mathcal{L}_q^{\mathrm{gen}} = \left\{ l \;\middle|\; 1 \leq l < \frac{1}{2}l_{\mathrm{opt}}(y_q) \text{ or } \frac{3}{2}l_{\mathrm{opt}}(y_q) < l < q/2 \right\}. \tag{E139}$$

It is now straightforward to verify that

$$\sqrt{\frac{N}{l(q-2l)(N-q+l)}} \leq \sqrt{\frac{4N}{l_{\mathrm{opt}}(q-3l_{\mathrm{opt}})l_{op}}} \text{ for } l \in \mathcal{L}_q^{\max}, \tag{E140}$$

where for clarity we surpass the dependence of l_{opt} on q. Moreover, from Eqs. (E130) and (E131) we get

$$NG(x_l, y_q) \leq NG_{\mathrm{opt}}(y_q) - \frac{(l-l_{\mathrm{opt}})^2}{3l_{\mathrm{opt}}} \text{ for } l \in \mathcal{L}_q^{\max}, \tag{E141}$$

$$NG(x_l, y_q) \leq NG_{\mathrm{opt}}(y_q) - 2x_{\mathrm{opt}}^2 N \text{ for } l \in \mathcal{L}_q^{\mathrm{gen}}. \tag{E142}$$

Finally, we arrive at the following bound:

$$\frac{\sqrt{2}CA}{c}\sum_{q>0.925N}^{N}\sum_{l=1}^{l<q/2}\sqrt{\frac{N}{l(q-2l)(N-q+l)}}\exp[NG(x_l,y_q)] \tag{E143}$$

$$\leq \frac{\sqrt{2}CA}{c}\sum_{q>0.925N}^{N}\exp[NG_{\mathrm{opt}}(y_q)]\sqrt{\frac{4N}{l_{\mathrm{opt}}(q-3l_{\mathrm{opt}})l_{\mathrm{opt}}}}\sum_{l\in\mathcal{L}_q^{\max}}\exp\left(-\frac{(l-l_{\mathrm{opt}})^2}{3l_{\mathrm{opt}}}\right) \tag{E144}$$

$$+\sqrt{\frac{N}{N-2}}\frac{CA}{c}\sum_{q>0.925N}^{N} q\exp[NG_{\mathrm{opt}}(y_q)]\exp\left(-2x_{\mathrm{opt}}^2 N\right), \tag{E145}$$

where we use Eq. (E132) to get Eq. (E145). We first analyze the second sum. By using Eq. (E125) we obtain

$$\sum_{q>0.925N}^{N} q\exp[NG_{\rm opt}(y_q)] \le N \sum_{q>0.925N}^{N} \exp\left[\frac{(N-q)^2}{N}\right] \le N\left(1+\frac{\sqrt{\pi N}}{2}\right) \le N^{\frac{3}{2}}, \tag{E146}$$

where we use

$$\sum_{x=0}^{\infty}\exp\left(-\frac{x^2}{a}\right) \le 1+\int_0^{\infty} dx\exp\left(-\frac{x^2}{a}\right) = 1+\frac{\sqrt{\pi a}}{2}, \tag{E147}$$

valid for all $a>0$, and $N \ge 100$. Importantly, for $q>0.925N$ (which corresponds to $y \ge 0.925$), we have $x_{\rm opt} \ge 0.03$. Using this and assuming $N \ge 7000$, we finally obtain

$$\sqrt{\frac{N}{N-2}\frac{CA}{c}}\sum_{q>0.925N}^{N} q\exp[NG_{\rm opt}(y_q)]\exp\left(-2x_{\rm opt}^2N\right) \le \sqrt{\frac{N}{N-2}\frac{CA}{c}}N^{\frac{3}{2}}\exp\left(-\frac{9}{5000}N\right) \le 1. \tag{E148}$$

We use similar methods to bound (E144). First, we upper bound the exponential sum

$$\sum_{l\in\mathcal{L}_q^{\max}}\exp\left(-\frac{(l-l_{\rm opt})^2}{3l_{\rm opt}}\right) \le 1+\sqrt{\pi 3l_{\rm opt}} \le \frac{10}{3}\sqrt{l_{\rm opt}}, \tag{E149}$$

which allows estimate

$$\sqrt{\frac{4N}{l_{\rm opt}(q-3l_{\rm opt})l_{\rm opt}}}\sum_{l\in\mathcal{L}_q^{\max}}\exp\left(-\frac{(l-l_{\rm opt})^2}{3l_{\rm opt}}\right) \le \frac{10}{3}\sqrt{\frac{4N}{l_{\rm opt}(q-3l_{\rm opt})}} \le \frac{10}{3}\sqrt{\frac{4}{(0.03)(0.7N)}} = \frac{20}{3}\sqrt{\frac{1000}{21N}}, \tag{E150}$$

where in the second inequality we use that for $q \ge 0.925N$ we have $l_{\rm opt}(y_q) \ge 0.03N$ and $q-3l_{\rm opt}(y_q) \ge 0.7N$. Inserting thin inequality to Eq. (E144) and again using Eq. (E146) gives that for $N \ge 7000$

$$\frac{\sqrt{2}CA}{c}\sum_{q>0.925N}^{N}\sum_{l=1}^{l<q/2}\sqrt{\frac{N}{l(q-2l)(N-q+l)}}\exp[NG(x_l,y_q)] \le 1+\frac{\sqrt{2}CA}{c}\sqrt{\frac{1000}{21}} \le 12.7. \tag{E151}$$

Combining this estimate with the bound (E137) and using Eq. (E133), we finally obtain that for $N \ge 7000$

$$\mathcal{B}_{\rm gen} \le 14.7. \tag{E152}$$

Finally, combining bounds (E117), (E121), and (E152) together with $\mathcal{B}_{k=0}=1$ in inequality (E107) we see that for $N \ge 7000$,

$$\begin{aligned}\mathrm{tr}(\mathbb{P}_{\rm act}\Psi_{\rm in}\otimes\Psi_{\rm in}) &\le \frac{1}{\sqrt{\pi N}}(\mathcal{B}_{q=0}+\mathcal{B}_{l=0}+\mathcal{B}_{q=2l}+\mathcal{B}_{\rm gen})\\ &\le \frac{16.2}{\sqrt{\pi N}}.\end{aligned} \tag{E153}$$

For $N \le 7000$, the validity of the bound can be verified numerically as shown in Fig. 7, which completes the proof. ■

APPENDIX F: EFFICIENT TOMOGRAPHY OF FLO UNITARIES

Here we prove Lemma 6, which establishes a bound concerning the stability of the active FLO representation, which is needed in the efficient tomographic scheme of Sec. IX.

Lemma 20: (Stability of active FLO representation.) *Consider two elements of the orthogonal group, $O, O' \in SO(2d)$, and let V and V' be the corresponding active FLO unitaries, i.e., $V=\Pi_{\rm act}(O)$ and $V'=\Pi_{\rm act}(O')$. Let Φ_V and $\Phi_{V'}$ be the unitary channels defined by V and V', respectively. Then, the following inequality is satisfied:*

$$\|\Phi_V-\Phi_{V'}\|_\diamond \le 2d\|O-O'\|. \tag{F1}$$

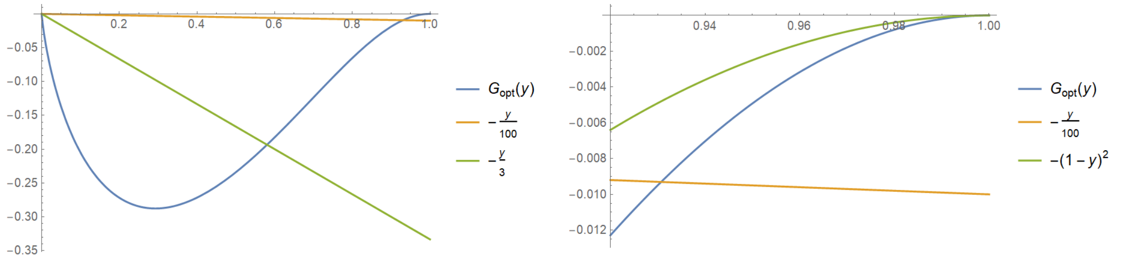


FIG. 13. Function $G_{\mathrm{opt}}(y) = G(x_{\mathrm{opt}}(y), y)$ where G is defined in Eq. (E123) and $x_{\mathrm{opt}}(y)$ is defined in Eq. (E124). The function is presented alongside simple analitical lower bounds are valid in specific intervals formulated in Eq. (E125).

Proof. The proof will rely on representation theoretic methods, however, as we have noted, Π_{act} is a projective representation of SO$(2d)$ and not a proper representation. Instead, we use $\Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}$, which is already a proper representation of SO$(2d)$. Thus, we bound the diamond norm difference between the unitary channels $\phi_{V\otimes V}$ and $\phi_{V'\otimes V'}$ corresponding to the unitaries $V \otimes V = \Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O)$ and $V' \otimes V' = \Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O')$, respectively, and then use the inequalities

$$\begin{aligned} \|\Phi_V - \Phi_{V'}\|_\diamond &\le \|\Phi_{V\otimes V} - \Phi_{V'\otimes V'}\|_\diamond \\ &\le 2\|V \otimes V - V' \otimes V'\|. \end{aligned} \tag{F2}$$

Here the first inequality follows directly from the definition of the diamond norm, while the second is a standard inequality relating the diamond norm distance of unitary channels to the operator norm distance of unitaries (see, e.g., Ref. [120]).

Thus, our proof strategy is to upper bound $\|V \otimes V - V' \otimes V'\| = \|\Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O) - \Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O')\|$. For this we use the decomposition of the $\Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}$ into subrepresentations in the following way [121]:

$$\Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}} \cong \bigoplus_{s=0}^{d-1} \left(\bigwedge^{s} \Pi \right)^{\oplus 2} \oplus \bigwedge^{d} \Pi, \tag{F3}$$

where Π denotes the defining representation of SO$(2d)$ and its ℓth antisymmetric tensor power $\bigwedge^{\ell} \Pi$ is given by

$$\bigwedge^{\ell} \Pi : \mathrm{SO}(2d) \longrightarrow \mathrm{U}\Big(\bigwedge^{\ell} (\mathbb{C}^{2d}) \Big), \tag{F4}$$

$$O \longmapsto O^{\otimes n}\big|_{\bigwedge^{\ell}(\mathbb{C}^{d})} . \tag{F5}$$

This decomposition immediately implies that

$$\begin{aligned} &\|V \otimes V - V' \otimes V'\| = \|\Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O) - \Pi_{\mathrm{act}} \otimes \Pi_{\mathrm{act}}(O')\| \\ &\le \max_{\ell \in [d]} \| \bigwedge^{\ell} \Pi(O) - \bigwedge^{\ell} \Pi(O')\| \le \max_{\ell \in [d]} \|O^{\otimes \ell} - O'^{\otimes \ell}\|. \end{aligned} \tag{F6}$$

Inserting the above inequality into Eq. (F2) and using that $\|O^{\otimes \ell} - O'^{\ell}\| \le \ell \|O - O'\|$ (and $\ell \le d$), we obtain

$$\|\Phi_V - \Phi_{V'}\|_\diamond \le 2d\|O - O'\|. \tag{F7}$$

■

APPENDIX G: #*P*-HARDNESS OF PROBABILITIES IN SHALLOW-DEPTH ACTIVE FLO CIRCUITS

We argued in Sec. VI that amplitudes of active FLO circuits are #*P*-hard to compute. Here we show that similarly strong simulation (i.e., computing output probabilities) of constant-depth active FLO circuits is hard. It has been proven in previous work [11] that under certain conditions, nonuniversal circuit families of shallow depth are hard to simulate under plausible conjectures, which in addition implies that the output probabilities are #*P*-hard. Specifically, it is required that the postselected version of the circuit family is universal for quantum computation. This method is not robust as it shows only that exactly computing the output probabilities are hard, nonetheless it may be of interest that such hardness results can be obtained for constant-depth active FLO circuits. The required theorem is as follows.

Theorem 9: *Let $\mathcal{F}$ be a restricted family of quantum circuits. If circuits from $\mathcal{F}$ with the added power of postselection can simulate the output probability distributions of universal quantum circuits with postselection (i.e., $\mathcal{F}$ is universal with postselection) then computing the output probabilities (strong simulation) of circuits in $\mathcal{F}$ is #P-hard.*

Proof. Similar results have been proven in Refs. [10,11] and later in other works related to active FLO [77]. Let C be some circuit with gates from a universal gate set and let $P_C(\mathbf{y})$ be the output probability of result $\mathbf{y}$. By hypothesis, with the power of postselection we can use a circuit F from $\mathcal{F}$ to simulate C and thus $P_C(\mathbf{y}) = P_F(\mathbf{y}_* | 00\cdots 0) =$

$P_F(\mathbf{y}_*00\cdots0)/P_F(00\cdots0)$, where $\mathbf{y}_*$ is potentially a bitstring encoding $\mathbf{y}$ (which will be our case below). This directly implies that if we could compute the output probabilities of F then this would allow for computing the output probabilities of C. Since universal circuits are known to include #P-hard instances, the result follows. ∎

In what follows, we always assume that the active FLO circuits are supplied with auxiliary states $|\Psi_4\rangle$. Throughout this section we consider the encoding $|0_L\rangle = |00\rangle$ and $|1_L\rangle = |11\rangle$. To prove that computing the probabilities of shallow-depth active FLO circuits is #P-hard, we prove now Lemma 19.

Lemma 21: *Constant-depth active FLO circuits supplied with auxiliary states $|\Psi_4\rangle$ with the added power of postselection are universal.*

To prove this, we follow Ref. [51], which showed similar results in the context of boson sampling. The starting point is the brickwork graph state, which allows for universal computation on the measurement-based quantum computation (MBQC) scheme. We can write the preparation of the brickwork graph state plus measurements on the state as a single circuit with adaptive measurements. If we are given the power to postselect measurements, then the preparation of the graph state requires a constant depth circuit with single-qubit gates and CZ gates. If we can simulate these gates with constant-depth active FLO circuits and postselection, then this would imply Lemma 19. Using the encoding defined above, we show Theorem 10, which directly implies Lemma 19.

Theorem 10: *Active FLO acting on an initial state consisting of tensor products of $|\Psi_4\rangle$ with the added power of postselection can simulate single-qubit gates and CZ with constant-depth circuits. These simulations are at the logical level using the encoding above.*

Proof. As explained before, the circuit induced by the brickwork state with postselection is universal and of constant depth, consisting of single qubit gates and CZ gates. Using the encoding above we can simulate single-qubit gates and CZ gates in constant depth, then we can simulate the whole universal constant-depth circuit with a circuit from $\mathcal{C}_{\text{act}}$ and postselection.

That single-qubit gates at the logical level can be implemented with this encoding is already known [55]. Implementing CZ at the logical level will require the use of postselection and the auxiliary states $|\Psi_4\rangle$. First, we note that the state $|\Psi_4\rangle$ can be transformed into the state $|a_8\rangle = 1/\sqrt{2}(|0000\rangle + |1111\rangle)$ using only active FLO operations. This was shown previously in the proof of Lemma 15. Second, in Lemma 1 of Ref. [58] it is shown that using a single copy of $|a_8\rangle$ and particle-number measurements it is possible to implement a CZ at the logical level using the same encoding we use here. This two facts together imply that CZ can be implemented with active FLO circuits supplied by $|\Psi_4\rangle$ states and postselection. The auxiliary states can be swapped to the desired position when implementing a gate without incurring on extra negative signs with our encoding since the auxiliary states used are fermionic as, for example, argued in Ref. [59]. ∎

[1] C. Gidney and M. Ekerå, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits, (2019), arXiv e-prints, ArXiv:1905.09749.

[2] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. Girvin, L. Jiang, and *et*, Extending the lifetime of a quantum bit with error correction in superconducting circuits, Nature **536**, 441 (2016).

[3] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, and *et*, Fault-Tolerant Operation of a Quantum Error-Correction Code, (2020), arXiv e-prints, ArXiv:2009.11482.

[4] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[5] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multiport interferometers, Optica **3**, 1460 (2016).

[6] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental Realization of any Discrete Unitary Operator, Phys. Rev. Lett. **73**, 58 (1994).

[7] A. P. Lund, M. J. Bremner, and T. C. Ralph, Quantum sampling problems, bosonsampling and quantum supremacy, npj Quantum Inf. **3**, 15 (2017).

[8] A. W. Harrow and A. Montanaro, Quantum computational supremacy, Nature (London) 549, 203 (2017). ArXiv:1809.07442

[9] S. Bravyi, D. Gosset, and R. König, Quantum advantage with shallow circuits, Science **362**, 308 (2018).

[10] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory Comput. **4**, 143 (2013).

[11] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy, Proc. R. Soc. A **467**, 459 (2011).

[12] M. Bremner, A. Montanaro, and D. Shepherd, Average-Case Complexity versus Approximate Simulation of Commuting Quantum Computations, Phys. Rev. Lett. **117**, 080501 (2016).

[13] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. 14, 595 (2018). ArXiv:1608.00263

[14] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2019).

[15] R. Movassagh, Quantum supremacy and random circuits, (2019), arXiv e-prints, ArXiv:1909.06210.

[16] B. Fefferman and C. Umans, in *11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 61, edited by A. Broadbent (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2016), p. 1:1.

[17] T. Morimae, Hardness of classically sampling the one-clean-qubit model with constant total variation distance error, Phys. Rev. A **96**, 040302 (2017).

[18] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, Architectures for Quantum Simulation Showing a Quantum Speedup, Phys. Rev. X 8, 021010 (2018). ArXiv:1703.00466

[19] A. Bouland, J. F. Fitzsimons, and D. E. Koh, in *33rd Computational Complexity Conference (CCC 2018)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 102, edited by R. A. Servedio (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2018), p. 21:1, iSSN: 1868-8969.

[20] H. Pashayan, S. D. Bartlett, and D. Gross, From estimation of quantum probabilities to simulation of quantum circuits, Quantum **4**, 223 (2020).

[21] M. Yoganathan, R. Jozsa, and S. Strelchuk, Quantum advantage of unitary clifford circuits with magic state inputs, Proc. R. Soc. A **475**, 20180427 (2019).

[22] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Gaussian Boson Sampling, Phys. Rev. Lett. **119**, 170501 (2017).

[23] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, Boson Sampling from a Gaussian State, Phys. Rev. Lett. **113**, 100502 (2014).

[24] J. Haferkamp, D. Hangleiter, A. Bouland, B. Fefferman, J. Eisert, and J. Bermejo-Vega, Closing gaps of a quantum advantage with short-time Hamiltonian dynamics, (2019), arXiv e-prints, ArXiv:1908.08069.

[25] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, and *et*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[26] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, Anticoncentration theorems for schemes showing a quantum speedup, Quantum **2**, 65 (2018).

[27] A. Harrow and S. Mehraban, Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates, (2018), arXiv e-prints, ArXiv:1809.06957.

[28] D. J. Brod, E. F. Galvão, A. Crespi, R. Osellame, N. Spagnolo, and F. Sciarrino, Photonic implementation of boson sampling: A review, Adv. Photonics **1**, 034001 (2019).

[29] H. Wang, J. Qin, X. Ding, M.-C. Chen, S. Chen, X. You, Y.-M. He, X. Jiang, L. You, Z. Wang, and *et*, Boson Sampling with 20 Input Photons and a 60-Mode Interferometer in a 10^{14}-Dimensional Hilbert Space, Phys. Rev. Lett. **123**, 250503 (2019).

[30] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, and *et*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[31] A. Bouland, B. Fefferman, Z. Landau, and Y. Liu, Noise and the frontier of quantum supremacy, (2021), arXiv e-prints, ArXiv:2102.01738.

[32] A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, J. Eisert, and *et*, Quantum computational supremacy via high-dimensional Gaussian boson sampling, (2021), arXiv preprint, ArXiv:2102.12474.

[33] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, Phys. Rev. A **65**, 032325 (2002).

[34] E. Knill, Fermionic Linear Optics and Matchgates, (2001), arXiv e-prints, ArXiv:quant-ph/0108033.

[35] L. Valiant, Quantum circuits that can be simulated classically in polynomial time, SIAM J. Comput. **31**, 1229 (2002).

[36] E. Bocquillon, V. Freulon, F. D. Parmentier, J.-M. Berroir, B. Plaçais, C. Wahl, J. Rech, T. Jonckheere, T. Martin, C. Grenier, and *et*, Electron quantum optics in ballistic chiral conductors, Ann. Phys. **526**, 1 (2014).

[37] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, and *et*, Hartree-Fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).

[38] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, Phys. Rev. Lett. **120**, 110501 (2018).

[39] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions, Phys. Rev. Appl. **9**, 044036 (2018).

[40] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer, Quantum Sci. Technol. **4**, 045005 (2019).

[41] B. Foxen, C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Z. Chen, K. Satzinger, R. Barends, and *et*, Demonstrating a Continuous Set of Two-Qubit Gates for Near-Term Quantum Algorithms, Phys. Rev. Lett. **125**, 120504 (2020).

[42] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anti-concentrate in log depth, (2020), arXiv e-prints, ArXiv:2011.12277.

[43] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **291**, 257 (2009).

[44] Y. Wu, *et al.*, Strong quantum computational advantage using a superconducting quantum processor, (2021), arXiv e-prints, ArXiv:2106.14734.

[45] H.-S. Zhong, *et al.*, Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light, (2021), arXiv e-prints, ArXiv:2106.15534.

[46] D. Shepherd and M. J. Bremner, Temporally unstructured quantum computation, Proc. R. Soc. A: Math., Phys. Eng. Sci. **465**, 1413 (2009).

[47] Informally, #*P* is the complexity class of counting solutions to problems that can be efficiently verified. For more details, see Ref. [54].

[48] B. M. Terhal and D. P. DiVincenzo, Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games, Quantum Inf. Comput. **4**, 134 (2004).

[49] S. Fenner, F. Green, S. Homer, and R. Pruim, Determining acceptance possibility for a quantum computation is hard for the polynomial hierarchy, Proc. R. Soc. London. Ser. A: Math., Phys. Eng. Sci. **455**, 3953 (1999).

[50] E. Farhi and A. W. Harrow, Quantum Supremacy through the Quantum Approximate Optimization Algorithm, (2016), arXiv e-prints, ArXiv:1602.07674.

[51] D. J. Brod, Complexity of simulating constant-depth boson sampling, Phys. Rev. A **91**, 042316 (2015).

[52] X. Gao, S.-T. Wang, and L.-M. Duan, Quantum Supremacy for Simulating a Translation-Invariant Ising Spin Model, Phys. Rev. Lett. **118**, 040502 (2017). publisher: American Physical Society

[53] L. Stockmeyer, On approximation algorithms for #P, SIAM J. Comput. **14**, 849 (1985).

[54] S. Arora and B. Barak, *Computational Complexity: A Modern Approach* (Cambridge University Press, USA, 2009), 1st ed.

[55] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, Ann. Phys. (N. Y) **298**, 210 (2002).

[56] D. A. Ivanov, Computational complexity of exterior products and multiparticle amplitudes of noninteracting fermions in entangled states, Phys. Rev. A **96**, 012322 (2017).

[57] D. A. Ivanov and L. Gurvits, Complexity of full counting statistics of free quantum particles in product states, Phys. Rev. A **101**, 012303 (2020).

[58] S. Bravyi, Universal quantum computation with the $\nu = 5/2$ fractional quantum Hall state, Phys. Rev. A **73**, 042313 (2006).

[59] M. Hebenstreit, R. Jozsa, B. Kraus, S. Strelchuk, and M. Yoganathan, All Pure Fermionic Non-Gaussian States are Magic States for Matchgate Computations, Phys. Rev. Lett. **123**, 080503 (2019).

[60] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, Sufficient Conditions for Efficient Classical Simulation of Quantum Optics, Phys. Rev. X **6**, 021039 (2016).

[61] D. Hangleiter, M. Kliesch, J. Eisert, and C. Gogolin, Sample Complexity of Device-Independently Certified "Quantum Supremacy", Phys. Rev. Lett. **122**, 210502 (2019).

[62] S. Aaronson and A. Arkhipov, Boson sampling is far from uniform, Quantum Inf. Comput. **14**, 1383 (2014).

[63] U. Chabaud, F. Grosshans, E. Kashefi, and D. Markham, Efficient verification of boson sampling, (2020), arXiv preprint ArXiv:2006.03520.

[64] M. Oszmaniec and Z. Zimborás, Universal Extensions of Restricted Classes of Quantum Operations, Phys. Rev. Lett. 119, 220502 (2017). ArXiv:1705.11188

[65] S. Bravyi, Lagrangian representation for fermionic linear optics, Quantum Info. Comput. **5**, 216 (2005).

[66] S. Bravyi and R. Koenig, Classical simulation of dissipative fermionic linear optics, (2011), arXiv e-prints, ArXiv:1112.2184.

[67] R. Jozsa and A. Miyake, Matchgates and classical simulation of quantum circuits, Proc. R. Soc. A: Math., Phys. Eng. Sci. **464**, 3089 (2008).

[68] D. J. Brod and E. F. Galvão, Geometries for universal quantum computation with matchgates, Phys. Rev. A **86**, 052307 (2012).

[69] D. J. Brod, Efficient classical simulation of matchgate circuits with generalized inputs and measurements, Phys. Rev. A **93**, 062332 (2016).

[70] S. Bravyi and A. Kitaev, Universal quantum computation with ideal clifford gates and noisy ancillas, Phys. Rev. A **71**, 022316 (2005).

[71] S. Bravyi, Classical capacity of fermionic product channels, (2005), arXiv e-prints, ArXiv:quant-ph/0507282.

[72] F. de Melo, P. Ćwikliński, and B. M. Terhal, The power of noisy fermionic quantum computation, New J. Phys. **15**, 013015 (2013).

[73] M. Oszmaniec, J. Gutt, and M. Kuś, Classical simulation of fermionic linear optics augmented with noisy ancillas, Phys. Rev. A **90**, 020302 (2014).

[74] D. J. Brod and E. F. Galvão, Extending matchgates into universal quantum computation, Phys. Rev. A **84**, 022310 (2011).

[75] Z. Zimborás, R. Zeier, M. Keyl, and T. Schulte-Herbrüggen, A dynamic systems approach to fermions and their relation to spins, EPJ Quantum Technol. **1**, 11 (2014).

[76] C. Beenakker, D. DiVincenzo, C. Emary, and M. Kindermann, Charge Detection Enables Free-Electron Quantum Computation, Phys. Rev. Lett. **93**, 020501 (2004).

[77] M. Hebenstreit, R. Jozsa, B. Kraus, and S. Strelchuk, Computational power of matchgates with supplementary resources, (2020), arXiv e-prints, ArXiv:2007.08231.

[78] A. Arkhipov, Boson sampling is robust against small errors in the network matrix, Phys. Rev. A **92**, 062326 (2015).

[79] The FLO operators themselves form a group isomorphic to the universal cover of SO$(2d)$, called spin$(2d)$.

[80] R. W. Goodman and N. R. Wallach, *Symmetry, Representations, and Invariants*, Graduate Texts in Mathematics (Springer, Dordrecht, 2009).

[81] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction* (Springer, Basel, Switzerland, 2000).

[82] Y. Kondo, R. Mori, and R. Movassagh, Improved robustness of quantum supremacy for random circuit sampling, (2021), arXiv e-prints, ArXiv:2102.01960.

[83] M. Lobino, D. Korystov, C. Kupchak, E. Figueroa, B. C. Sanders, and A. I. Lvovsky, Complete characterization of quantum-optical processes, Science **322**, 563 (2008).

[84] S. Rahimi-Keshari, A. Scherer, A. Mann, A. T. Rezakhani, A. I. Lvovsky, and B. C. Sanders, Quantum process tomography with coherent states, New J. Phys. **13**, 013006 (2011).

[85] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White, Direct characterization of linear-optical networks, Opt. Express **21**, 13450 (2013).

[86] N. Schuch and J. Siewert, Natural two-qubit gate for quantum computation using the xy interaction, Phys. Rev. A **67**, 032301 (2003).

[87] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, and Bucher, Qiskit: An open-source framework for quantum computing (2019).

[88] J. M. Arrazola and T. R. Bromley, Using Gaussian Boson Sampling to Find Dense Subgraphs, Phys. Rev. Lett. **121**, 030503 (2018).

[89] J. M. Arrazola, T. R. Bromley, and P. Rebentrost, Quantum approximate optimization with gaussian boson sampling, Phys. Rev. A **98**, 012322 (2018).

[90] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, Boson sampling for molecular vibronic spectra, Nat. Photonics **9**, 615 (2015).

[91] J. Huh and M.-H. Yung, Vibronic boson sampling: Generalized gaussian boson sampling for molecular vibronic spectra at finite temperature, Sci. Rep. **7**, 1 (2017).

[92] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, Molecular docking with gaussian boson sampling, Sci. Adv. **6**, eaax1950 (2020).

[93] M. Schuld, K. Brádler, R. Israel, D. Su, and B. Gupt, Measuring the similarity of graphs with a gaussian boson sampler, Phys. Rev. A **101**, 032314 (2020).

[94] H. Ikai, On the theory of pfaffians based on exponential maps in exterior algebras, Linear Algebra Appl. **434**, 1094 (2011).

[95] A. E. Moylett and P. S. Turner, Quantum simulation of partially distinguishable boson sampling, Phys. Rev. A **97**, 062329 (2018).

[96] J. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandao, and A. W. Harrow, Efficient classical simulation of random shallow 2D quantum circuits, (2019), arXiv e-prints, ArXiv:2001.00021.

[97] J. Helsen, S. Nezami, M. Reagor, and M. Walter, Matchgate benchmarking: Scalable benchmarking of a continuous family of many-qubit gates, (2020), arXiv preprint ArXiv:2011.13048.

[98] To simplify expressions for active FLO, we use the bound $\binom{4N}{2N} \geq 2^{8N}/\sqrt{\pi 4N}$.

[99] W. Lichtenstein, A system of quadrics describing the orbit of the highest weight vector, Proc. Am. Math. Soc. **84**, 605 (1982).

[100] M. Kuś and I. Bengtsson, "Classical" quantum states, Phys. Rev. A **80**, 022319 (2009).

[101] M. Oszmaniec and M. Kuś, Universal framework for entanglement detection, Phys. Rev. A **88**, 052328 (2013).

[102] M. Oszmaniec and M. Kuś, Fraction of isospectral states exhibiting quantum correlations, Phys. Rev. A **90**, 010302 (2014).

[103] M. Oszmaniec, Applications of differential geometry and representation theory to description of quantum correlations, (2014), arXiv e-prints, ArXiv:1412.4657.

[104] C. Lautemann, Bpp and the polynomial hierarchy, Inf. Process. Lett. **17**, 215 (1983).

[105] S. Toda, PP is as hard as the polynomial-time hierarchy, SIAM J. Comput. **20**, 865 (1991).

[106] R. Goodman and N. R. Wallach, *Symmetry, Representations, and Invariants* (Springer, New York, 2009).

[107] G. Aubrun and S. J. Szarek, *Alice and Bob meet Banach: the interface of asymptotic geometric analysis and quantum information theory*, Mathematical surveys and monographs (American Mathematical Society, Providence, RI, 2017).

[108] R. Paturi, in *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, STOC '92 (Association for Computing Machinery, New York, NY, USA, 1992), p. 468.

[109] D. Coppersmith and T. J. Rivlin, The growth of polynomials bounded at equally spaced points, SIAM J. Math. Anal. **23**, 970 (1992).

[110] E. A. Rakhmanov, Bounds for polynomials with a unit discrete norm, Ann. Math. **165**, 55 (2007).

[111] M. E. Dyer, L. A. Goldberg, C. S. Greenhill, and M. Jerrum, in *Proceedings of the Third International Workshop on Approximation Algorithms for Combinatorial Optimization*, APPROX '00 (Springer-Verlag, Berlin, Heidelberg, 2000), p. 108.

[112] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking, Nat. Rev. Phys. **2**, 382 (2020).

[113] R. Mathias, Perturbation bounds for the polar decomposition, SIAM J. Matrix Anal. Appl. **14**, 588 (1993).

[114] J. A. Tropp, An introduction to matrix concentration inequalities, Found. Trends Mach. Learn. **8**, 1 (2015).

[115] E. M. Rains, Polynomial invariants of quantum codes, IEEE Trans. Inf. Theory **46**, 54 (2000).

[116] H. Elvang, P. Cvitanović, and A. D. Kennedy, Diagrammatic young projection operators for u(n), J. Math. Phys. **46**, 043501 (2005).

[117] A. J. Coleman and V. I. Yukalov, *Reduced Density Matrices: Coulson's Challenge* Vol. 72 (Springer Science & Business Media, Berlin, Heidelberg, 2000).

[118] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes* (North-Holland Pub. Co., Oxford, England, 1983).

[119] It is easy to check that $3/2x_{\mathrm{opt}}(y) \leq y/2$.

[120] M. Oszmaniec, A. Sawicki, and M. Horodecki, Epsilon-nets, unitary designs and random quantum circuits, IEEE Trans. Inf. Theory, **68**, 989 (2022).

[121] J. Fuchs and C. Schweigert, *Symmetries, Lie Algebras and Representations: A Graduate Course for Physicists* (Cambridge University Press, Cambridge, England, 2003).