Computing with Highly Mixed States

ANDRIS AMBAINIS
University of Waterloo, Waterloo, Canada, Ont.

LEONARD J. SCHULMAN
California Institute of Technology, Pasadena, California

AND

UMESH VAZIRANI
University of California, Berkeley, California

Abstract. Device initialization is a difficult challenge in some proposed realizations of quantum computers, and as such, must be treated as a computational resource. The degree of initialization can be quantified by $k$, the number of clean qubits in the initial state of the register. In this article, we show that unless $m \in O(k + \log n)$, oblivious (gate-by-gate) simulation of an ideal $m$-qubit quantum circuit by an $n$-qubit circuit with $k$ clean qubits is impossible. Effectively, this indicates that there is no avoiding physical initialization of a quantity of qubits proportional to that required by the best ideal quantum circuit.

Categories and Subject Descriptors: F.1 [Theory of Computation]: Computation by Abstract Devices
General Terms: Algorithms, Theory
Additional Key Words and Phrases: Quantum Computation

1. Introduction

The field of quantum computation is driven by the recent realization that there is a profound difference between quantum and classical mechanics in the resources

A. Ambainis was supported in part by NSERC, ARDA, CIAR, and an IQC University Professorship.
L. J. Schulman was supported in part by National Science Foundation (NSF) grant CCF-0524828 and by the Institute for Quantum Information under NSF EIA-0086038, NSF PHY-0456720, and ARO W911NF-05-1-0294.
U. Vazirani was supported in part by NSF CCF-0524837 and ARO DAAD19-03-1-0082.

Authors’ addresses: A. Ambainis, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 5A3, Canada, e-mail: ambainis@math.uwaterloo.ca; L. J. Schulman, Department of Computer Science, California Institute of Technology, Pasadena, CA 91125, e-mail: schulman@caltech.edu; U. Vazirani, Computer Science Division, University of California, Berkeley, Berkeley, CA 94720, e-mail: vazirani@cs.berkeley.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 1515 Broadway, New York, NY 10036 USA, fax: +1 (212) 869-0481, or permissions@acm.org.
© 2006 ACM 0004-5411/06/0500-0507 $5.00

time, communication, interactive proof complexity, etc.) required for computation. The peculiar nature of quantum mechanics also introduces new computational considerations that have no classical counterpart. In this article, we study a computational resource—initialization—that is important in the setting of quantum computation, but which has little relevance in classical computation. Ideally, a quantum computation is a sequence of local unitary transformations applied to a register of qubits which are initially in the state $|0^n\rangle$, followed by a measurement. However, initializing the state of the quantum register can be quite challenging. This is currently the limiting resource for certain implementations, such as NMR quantum computing [Cory et al. 1997; Gershenfeld and Chuang 1997; Chuang et al. 1998; Vandersypen et al. 2000]. Initializations begins with a physical cooling (polarizing) stage, which results in each of the $n$ qubits independently being in state $|0\rangle$ with probability $1/2 + \epsilon/2$ and in state $|1\rangle$ with probability $1/2 - \epsilon/2$. Current implementations in the laboratory all rely on a second phase of initialization which creates a pseudopure state which is a noisy version of $|0^n\rangle$. Unfortunately and unavoidably in this method, the signal to noise ratio decreases exponentially in $n$ [Nielsen 1999].

A different scheme for this second phase of initialization via algorithmic cooling was proposed in Schulman and Vazirani [1999]. This scheme results in $m \approx \epsilon^2 n$ (almost) perfectly initialized qubits. However, currently achievable values of $\epsilon$ in liquid state NMR are around $10^{-5}$, thus requiring $n$ to be impractically large. Moreover, the bounds on $m$ in Schulman and Vazirani [1999] are tight to within lower order additive terms.

Is initialization to the state $|0^n\rangle$ really necessary for general quantum computation? This question was raised by Knill and Laflamme [1998], who showed how to perform an interesting computation on a very weakly initialized quantum computer. The model they formalized is the $k$ clean qubit model, which assumes that $k$ of the qubits are perfectly initialized in the state $|0^k\rangle$, and the remaining $n - k$ are in the maximally mixed state (i.e., each qubit independently and uniformly distributed over $\{|0\rangle, |1\rangle\}$). The algorithm in Schulman and Vazirani [1999] can transform $n$ qubits, which all start out with bias $\epsilon = \sqrt{k/n}$ into $k$ almost clean qubits and $n - k$ almost maximally mixed ones. Because of that, any positive result about the $k$-clean qubit model also implies a positive result about the NMR model in which each qubit independently has a bias of $\epsilon$.

Knill and Laflamme [1998] adapted Kitaev’s phase estimation technique to show how to use a $1$-clean qubit device to estimate the trace of an $n$-qubit quantum circuit. No efficient classical algorithm is known for this task. This raised the possibility that preparation of strings of clean qubits is not a necessary prerequisite for running quantum algorithms.

This article mostly dispels this possibility. With the exception of the Knill–Laflamme algorithm, all known quantum algorithms are designed for pure-state quantum computers. We introduce the model of oblivious simulations where a pure-state quantum circuit is simulated by a mixed-state quantum computer by simulating each elementary gate of the pure-state circuit and then composing the resulting circuits. When an oblivious simulation is possible, it enables us to simulate an arbitrary pure-state circuit. We then show that any oblivious simulation of pure-state quantum algorithms on mixed-state quantum computers requires that the number of clean qubits $k$ be almost as large as the number of qubits in the pure-state quantum algorithm (unless $n$ is exponentially large). Thus we show that absent special purpose designs for mixed state algorithms, general purpose quantum
computation in mixed state devices will hinge upon improvements in laboratory techniques for qubit cooling.

We begin on the positive side by showing that the one clean qubit model can be used to efficiently simulate any logarithmic depth classical circuit, thus solving any problem in $NC^1$ efficiently. This is somewhat surprising because an easy argument shows that in the absence of clean qubits ($k = 0$), no computation at all is possible.

However, the primary question is whether the one clean qubit model (or $k$ clean qubits for small $k$) can be used to simulate a general quantum circuit on $m$ qubits for $m$ substantially larger than $k$. Our main result is negative—we show that an oblivious simulation of an $m$ qubit quantum computation is possible in the $k$ clean qubit model only if $m \leq (2k + \log n)(1 + o(1))$. This is a consequence of the following (stronger) geometric statement which is our main theorem: suppose that a collection of $2^m$ subspaces of dimension $2^{n-k}$ in $\mathbb{C}^{2^n}$ have the property that the intersection of every pair has dimension less than $(1 - \frac{1}{\text{poly}(m)})2^{n-k}$. Then, unless $m$ satisfies the bound above, there is no collection of unitary operators on $\mathbb{C}^{2^n}$ that induces all possible permutations of the subspaces. The argument rests upon the representation theory of the symmetric group.

2. Preliminaries

If a physical system can be in $N$ distinguishable (classical) states, then its (pure) quantum state is a linear superposition of these $N$ states. Thus, a two state system (called a qubit) is represented by a unit vector in a two dimensional complex Hilbert space:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \quad \alpha, \beta \in \mathbb{C} \quad \text{and} \quad \alpha^2 + \beta^2 = 1.$$  

The state of an $n$ qubit system is described by a unit vector in a $2^n$ dimensional Hilbert space:

$$|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle.$$  

Measuring the state of the system in the standard basis yields the outcome $x$ with probability $|\alpha_x|^2$. Moreover, following the measurement, the new state of the system is $|x\rangle$. More generally, a measurement is specified by a set of projections $P_1, \ldots, P_m$ into orthogonal subspaces that span the underlying Hilbert space. The outcome of the measurement is $j$ with probability $|P_j|\langle \psi |\psi\rangle|^2$, and following the measurement, the new state of the system is $P_j |\psi\rangle$.

In actuality, there is generally some uncertainty about the state of the quantum system (due to entanglement with environment). This uncertainty is represented by a probability distribution over pure states. Such a system is referred to as a mixed state. There are, however, different distributions over pure states that cannot be distinguished by any physical measurement. A complete description of a mixed state consisting of pure states $|\psi_i\rangle$ with probabilities $p_i$ is its density matrix,

$$\sum_i p_i |\psi_i\rangle \langle \psi_i|.$$  

Two mixed states are physically distinguishable if and only if they have different density matrices.
For example, the mixture of $|0\rangle$ with probability $1/2$ and $|1\rangle$ with probability $1/2$ has density matrix $\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, as does the mixture of $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ with probability $1/2$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ with probability $1/2$.

When two mixed states have distinct density matrices (say $\rho_1$ and $\rho_2$), it is natural to ask how well they can be distinguished by a measurement. The answer is that the least-error measurement is one that measures in the basis that diagonalizes $\rho_1 - \rho_2$.

Diagonalizing the density matrix gives a mixture over orthogonal states. (Up to a unitary change of basis this can be regarded as a classical distribution). This is the mixture with minimum entropy consistent with the density matrix. The entropy of this distribution is invariant under reversible (unitary) quantum logic gates, and is referred to as the von Neumann entropy of the density matrix.

Quantum mechanics asserts that the evolution of a system is unitary. Each step amounts to multiplication of $|\psi\rangle$ by a certain unitary matrix $U$: $|\psi\rangle \rightarrow U|\psi\rangle$. In terms of density matrices, each computation step amounts to conjugation by $U$: $\rho \rightarrow U\rho U^{-1}$.

2.1. THE BASIC MODEL OF A QUANTUM COMPUTER. A quantum computer consists of the following:

- A quantum register which is initialized to some pure state $|\psi_0\rangle$, or more generally, to some density matrix $\rho_0$.
- The quantum algorithm is executed by performing a sequence of elementary gates on this quantum register. Each such gate is a unitary transformation that acts as the identity on all but two qubits. (As shown by Barenco et al. [1995], such gates are sufficient to implement an arbitrary unitary transformation.) The quantum computer is classically controlled, in two senses: first, the description of the sequence of elementary gates to be executed is produced using polynomial (classical) resources. Thus, their complexity is not limited in any way by the number of clean qubits. Second, a classical controller applies these gates to the quantum register without becoming entangled with it.
- Generally, it is important to allow for errors in the gates of the quantum computer. Since this article is primarily concerned with a lower bound, we allow error-free gates.
- The output of the algorithm is obtained by measuring one or more qubits of the quantum register in the standard basis. Of course, the whole process may be repeated a polynomial number of times.

As described above, a quantum computer consists of unitary transformations followed by one measurement at the end of computation. More generally, one could define quantum computation as an arbitrary sequence of unitary transformations and measurements, where a measurement can be followed by a unitary transformation which depends on the result of the measurement. In this article, we only consider the first, more restrictive definition, for two reasons:

- For the standard model of quantum computers, the two definitions are equivalent [Nielsen and Chuang 2000].
- For the $k$-clean qubit model they are not known to be equivalent. However, NMR quantum computers, which are the motivation for the $k$-clean qubit model, are
not capable of performing measurements in the middle of the computation. Thus, a sequence of unitary transformations, followed by a measurement at the end, is the right model for NMR quantum computing.

2.2. ONE QUBIT COMPUTATION. In this article, we consider the $k$-clean qubit model introduced by Knill and Laflamme [1998]. The starting state of a 1-clean qubit computation is a uniform distribution over all basis states in which the first bit is a $|0\rangle$; this is given by the density matrix

$$\rho_{\text{start}} = \left( \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right) \otimes \left( \begin{array}{cc} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{array} \right) \otimes \cdots \otimes \left( \begin{array}{cc} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{array} \right).$$

(1)

Equivalently, this is separable mixed state in which the first qubit is initialized to $|0\rangle$ while the rest are in states that are uniformly random ($|0\rangle$ with probability $1/2$ and $|1\rangle$ with probability $1/2$).

In the $k$-clean qubit model, $k$ qubits are initialized to $|0\rangle$ while the rest are in uniformly random states.

If we have a state where every qubit is slightly biased towards $|0\rangle$ (as in the previous section), one can apply the purification procedure by Schulman and Vazirani [1999] and transform it into a state where almost all the bias is concentrated on one qubit and the rest are almost completely random. Such a state is very close to the starting state for one slightly biased qubit computation. Therefore, any positive result about our model gives a corresponding positive result about NMR quantum computing.

We note that this is the weakest possible model of this type in which one can still do nontrivial computations. If all $n$ qubits are completely random, no computation is possible. Indeed, in this case, the start state is

$$\left( \begin{array}{cc} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{array} \right) \otimes \cdots \otimes \left( \begin{array}{cc} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{array} \right) = 2^{-n} I$$

where $I$ is the $2^n \times 2^n$ identity matrix. Applying any computation $U$ yields

$$U 2^{-n} I U^\dagger = 2^{-n} U U^\dagger = 2^{-n} I.$$

The final density matrix is therefore independent of the computation.

2.3. SUBSPACE INTERPRETATION. A pure state is a vector $|\psi\rangle = \sum_{x_1, \ldots, x_n} \alpha_{x_1, \ldots, x_n} |x_1 \cdots x_n\rangle$ in the Hilbert space $\mathbb{C}^{2^n}$. Similarly, the state of a $k$-clean qubit computer is described by a $2^{n-k}$-dimensional subspace of $\mathbb{C}^{2^n}$. For instance, the initial state $\rho_{\text{start}}$ in a one-clean qubit computer is described by the subspace $V_{\text{start}}$ that is spanned by all basis states with first bit $= 0$. The identification is justified by the fact that the same $\rho_{\text{start}}$ is defined by the uniform mixture of any orthonormal basis for the subspace. Applying $U$ to this state transforms the density matrix to $U \rho_{\text{start}} U^\dagger$, and the subspace $V_{\text{start}}$ to $UV_{\text{start}}$.

This correspondence will be extensively used in the proof of our lower bound on one-qubit computation (Sections 6–8). For brevity, we will often refer to subspaces as the states of computation.
2.4. REPRESENTATION THEORY. We now introduce the basic notions of representation theory needed to prove our main theorem. For more information on group representations, see Fulton and Harris [1991] and Sagan [2001].

**Representation.** A representation $\rho$ of a group $G$ is a homomorphism $\rho$ from $G$ to the group of linear transformations $GL(V)$ of a vector space $V$. This means that, for any $g, h \in G$, $\rho(gh) = \rho(g)\rho(h)$. If the mapping $\rho$ is clear from the context, we often call the space $V$ itself a representation of $G$.

**Irreducibility.** We say that a subspace $W$ is an invariant subspace of a representation $\rho$ if $\rho(g)W \subseteq W$ for all $g \in G$. The zero subspace and the subspace $V$ are always invariant. If no nonzero proper subspaces are invariant, the representation is said to be irreducible.

**Isomorphism.** Two representations $\rho : G \rightarrow GL(V)$ and $\rho' : G \rightarrow GL(W)$ are isomorphic if there is a bijective linear map $\varphi : V \rightarrow W$ such that $\varphi\rho(g) = \rho'(g)\varphi$ for any $g \in G$.

In this article, we mostly confine ourselves to finite groups and their representations over complex vector spaces. Under these conditions, we have the following facts:

**Unitarity.** Every representation is conjugate to a unitary representation $\rho$, one in which every $\rho(g)$ is unitary. (So henceforth all representations will be assumed to be unitary.)

**Irreps.** $G$ has finitely many nonisomorphic irreducible representations (irreps), each of finite dimension. Every group has the trivial representation $1$ of dimension $1$, which maps all group elements to the number 1.

**Complete Reducibility.** Every representation $V$ can be decomposed into irreps $W_i$, possibly with multiplicity: $V = W_1 \oplus \cdots \oplus W_k$. This decomposition is unique when written as: $V = \sum W_i \otimes E_{j_i}$, where the $W_i$ are nonisomorphic irreps, and $E_i = \oplus_{j=1}^{n_i} 1$. Here $j_i$ is just the multiplicity with which the irrep $W_i$ appears.

**Irreducible representations of $S_M$.** In this article, we rely on the representation theory of the symmetric group $S_M$. The irreducible representations of $S_M$ are in one-to-one correspondence with the partitions of $n$. A partition of $M$ is a sequence $(\lambda_1, \ldots, \lambda_k)$ of positive integers, with $\lambda_1 \geq \cdots \geq \lambda_k$ for which $\sum \lambda_i = M$. It is customary to identify the partition $\lambda = (\lambda_1, \ldots, \lambda_k)$ with a diagram consisting of $k$ rows of boxes, the $i$th row containing $\lambda_i$ boxes. We will let $\lambda$ stand for both the partition and the associated diagram. For example, the diagram corresponding to the partition $\lambda = (4, 4, 2, 1)$ is shown in Figure 1.

The irreducible representation associated with $\lambda$ is denoted $\rho_{\lambda}$. There is an explicit formula for the dimension of $\rho_{\lambda}$. This involves the notion of a hook: for a cell $(i, j)$ of a Young tableau $\lambda$, the $(i, j)$-hook $H_{i,j}$ is the collection of all cells of $\lambda$ which are

![Figure 1. The Young diagram of $\lambda = (4, 4, 2, 1)$.](image)
Computing with Highly Mixed States

FIG. 2. The hook-lengths for (4, 4, 2, 1).

beneath \((i, j)\) (but in the same column) or to the right of \((i, j)\) (but in the same row), including the cell \((i, j)\). The \textit{length} of the hook is the number of cells appearing in the hook. We use \(h_{i,j}\) to denote the length of \(H_{i,j}\). With this notation, the dimension of \(\rho_\lambda\) may be expressed:

\[
\dim \rho_\lambda = \frac{n!}{\prod_{i,j} h_{i,j}},
\]

this product being taken over all hooks \(h\) of \(\lambda\). Figure 2 shows the hook lengths for the partition \(\lambda = (4, 4, 2, 1)\). Formula (2) implies that the dimension of corresponding representation is

\[
\frac{11!}{7 \cdot 5 \cdot 3 \cdot 2 \cdot 6 \cdot 4 \cdot 2 \cdot 3} = 1320.
\]

\textit{Restriction.} A representation \(\rho\) of a group \(G\) is also automatically a representation of any subgroup \(H\). Note that even if a representation is irreducible over \(G\), it may no longer be irreducible when restricted to \(H\).

\textit{Restriction from} \(S_M\) \textit{to} \(S_{M-1}\). In particular, we will be considering the restrictions of irreducible representations of \(S_M\) to \(S_{M-1}\) obtained by fixing an element \(k \in \{1, \ldots , M\}\). Let \(\lambda\) be a partition of \(M\) and \(\rho_\lambda\) be the corresponding irreducible representation. Then, when we restrict to \(S_{M-1}\), \(\rho_\lambda\) decomposes into irreducible representations of \(S_{M-1}\) in the following way:

\[
\rho = \bigoplus_{\lambda^-} \rho_{\lambda^-}
\]

where \(\lambda^-\) ranges over all shapes of size \(M - 1\) that can be obtained by deleting an “inside corner” from \(\lambda\). (An inside corner is simply a point of the shape whose deletion leaves a legal shape.)

For example, the representation \(\rho_{\lambda}, \lambda = (4, 4, 2, 1)\) of \(S_{11}\) decomposes into three irreducible representations of \(S_{10}\). The Young diagrams of these representations are shown in Figure 3.

\textit{Compact Lie groups.} In Section 8, we will analyze representations of groups that can be infinite. A \textit{Lie group} is a group which is also a topological space, with the property that the group multiplication and group inverse are continuous maps on this topological space. A Lie group is \textit{compact} if it is compact as a topological space. For example, \(SU(N)\), the group of all unitary transformations on an \(N\)-dimensional space is a compact Lie group. Compact Lie groups have the complete reducibility property, described above [Adams 1982].

\textit{Schur’s Lemma.} If \(\rho\) and \(\rho'\) are two irreducible representations and \(\varphi\) is an isomorphism between them, then any other isomorphism \(\varphi'\) between \(\rho\) and \(\rho'\) is
c \varphi$ for some constant $c \in \mathbb{C}$. In particular, any isomorphism $\varphi$ from an irreducible representation $\rho$ to itself is a multiple of the identity map on $\rho$. Schur’s lemma is true both for finite groups and for compact Lie groups.

3. Positive Result: NC$^1$

We show that there is a natural classical complexity class, NC$^1$, that can be computed in the one-clean-bit model using just three qubits. The fact that such nontrivial computations can be performed by one-clean-qubit devices underscores the difficulty of bounding the power of this class.

3.1. NC$^1$, Branching Programs and One-Clean-Qubit Computation. NC$^1$ is the class of functions computable by fanin-2 Boolean circuits of polynomial size and logarithmic depth. By Barrington’s theorem [Barrington 1989], this class can be characterized by constant width permutation branching programs.

A branching program (BP) is a computing device that can be in one of finitely many states. It is specified by a sequence of instructions of the form $(i, \pi_0, \pi_1)$ where $i$ is an index for a variable $x_i$ and $\pi_0$ and $\pi_1$ are permutations on the states. If $x_i = 0$, $\pi_0$ is applied. If $x_i = 1$, $\pi_1$ is applied. The length of a BP is the number of instructions in it.

For any given boolean inputs $x_1, \ldots, x_n$, running a BP induces the permutation $\pi_{x_n} \cdots \pi_{x_1}$ on the states. We say that a BP $\omega$-computes a function $f(x_1, \ldots, x_n)$ if $\pi_{x_n} \cdots \pi_{x_1}$ is the identity when $f(x_1, \ldots, x_n) = 0$ and $\omega$ when $f = 1$.

**Theorem 1** [Barrington 1989]. Let $\omega$ be a five-cycle. NC$^1$ is equal to the class of languages $\omega$-recognized by a uniform family of width-5 polynomial length BPs.

**Proposition 1.** A three-qubit quantum computer initialized with one clean qubit can recognize every language in NC$^1$.

**Proof.** Let $L \in$ NC$^1$. Take the permutation branching program that $\omega$-recognizes $L$ where $\omega$ is the five-cycle $(000, 001, 010, 011, 100)$. We represent the five states of the permutation branching program by the states $|000\rangle$, $|001\rangle$, $|010\rangle$, $|011\rangle$, $|100\rangle$ of the three-qubit register.

The starting state of our one-clean qubit computation is the state in which the first qubit is $|0\rangle$ and all other qubits are uniformly random. This state corresponds to the subspace spanned by $|000\rangle$, $|001\rangle$, $|010\rangle$, $|011\rangle$. 

![Young diagrams of representations of $S_{10}$ contained in $\rho_\lambda$.](image)
We simulate the permutation branching program by a sequence of unitary transformations on 3 qubits. The transformation corresponding to a permutation \( \pi_i \) applies \( \pi_i \) to the basis states \(|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle \) and leaves the other basis states unchanged.

At the end, we measure the first qubit in the register. Two cases are possible:

1. \( x \notin L \). Then, the branching program maps each of the states \(|000\rangle, |001\rangle, |010\rangle, |011\rangle \) to itself. This means that the end state is the same as the starting state and measuring the first qubit always gives 0.

2. \( x \in L \). Then, the branching program permutes the basis states according to the transformation \( \omega \). This means that the end state is a probabilistic combination of \( \omega|000\rangle = |001\rangle, \omega|001\rangle = |010\rangle, \omega|010\rangle = |011\rangle, \omega|011\rangle = |100\rangle \). Therefore, measuring the first qubit of the end state yields a 1 with probability \( 1/4 \).

This result can be refined in two respects. First, instead of getting 1 with probability \( 1/4 \) in the second case, we can get it with probability 1. To show that, we first notice that the proof of Barrington’s theorem remains valid for eight-cycles instead five-cycles. Then, we take a BP that \( (000, 100, 001, 101, 010, 110, 011, 111) \)-recognizes \( L \) and apply the same argument. If \( x \in L \), then the mixture of states \(|000\rangle, |001\rangle, |010\rangle, |011\rangle \) is mapped to the mixture of \(|100\rangle, |101\rangle, |110\rangle, |111\rangle \) and measuring the first qubit gives 1 with probability 1.

Second, our result was extended by Ablayev et al. [2002] to show that the \( \text{NC}^1 \) simulation can be carried out on a computer with just one (clean) qubit.

3.2. \( \text{QNC}^1 \). It is illuminating to try to extend this simulation to \( \text{QNC}^1 \). \( \text{QNC}^1 \) is the class of languages recognized by quantum circuits of polynomial size and logarithmic depth [Moore and Nilsson 2001].

Notice that in Barrington’s procedure for simulating \( \text{NC}^1 \), each wire in the \( \text{NC}^1 \) circuit is simulated at some stage in the branching program. In the case of a \( \text{QNC}^1 \) circuit, the state of a wire is given by a qubit, which is, in general, entangled with the qubits carried by the other wires in the circuit. Therefore, the state of this wire cannot be expressed in isolation, and there appears to be no alternative to creating that entangled state as part of any simulation. Thus, the entire approach breaks down. One way to carry out such a construction, might be to apply a superposition of operations at each step: this extends the state space of the quantum computer and effectively provides many more clean qubits, making the model meaningless. Moreover, all proposed implementations of quantum computation involve a classical, time-varying sequence of operations, applied to a quantum register. Since the control is classical, in any oblivious simulation the entangled quantum state of the simulated circuit must be encoded within the quantum register.

4. Limits on Computability

4.1. REQUIREMENTS FOR \( k \)-CLEAN-QUBIT SIMULATION. We now turn to the main question of this article: Given a general \( m \)-qubit quantum circuit \( C \) that starts in the ideal \( |00 \ldots 0\rangle \) starting state, can we simulate it by an \( n \)-qubit circuit \( C' \) in the \( k \)-clean-qubit model? We seek an oblivious simulation that works for any circuit \( C \). In an oblivious simulation, we must encode the \( 2^m \) basis states of \( C \) as distinguishable states of the \( n \) qubit register of \( C' \). Recall that the general state of
$\mathcal{C}'$ is a density matrix corresponding to a subspace of dimension $2^{n-1}$ (in the one-clean qubit case) or $2^{n-k}$ (in the $k$-clean qubit case). Thus, we must demonstrate an encoding of $2^m$ basis states $|b\rangle$, $b \in \{0, 1\}^m$ by subspaces $A_b$ of dimension $2^{n-1}$ (or $2^{n-k}$). There are two requirements for this encoding.

1. **Distinguishability.** The states must be distinguishable in the following sense: since we can prepare several copies of any state by repeating the simulation, we only require that there be a sequence of measurements on $O(\text{poly}(n))$ many copies of the state, that (with high probability) uniquely identify the state. Indeed, it is possible to do this with $n = m$, as follows: take the subspaces spanned by the basis vectors in the sets $A_b = \{x \in \{0, 1\}^n : x \cdot b = 0 \mod 2\}$ for $b \in \{0, 1\}^m$.

Let $\rho_b$ be the mixed state over $A_b$. Measuring it gives a random vector $x$ such that $x \cdot b = 0$. Repeating this $n$ times gives $n$ such vectors $x_1, x_2, \ldots, x_n$. With high probability, there is only one $b$ such that $x_i \cdot b = 0$ for all $i$ and it can be found by solving the system of linear equations $x_i \cdot b = 0$ with standard methods. (This is similar to what happens in Simon’s quantum algorithm [Simon 1997].) These states are not, however, permutable.

Our impossibility result applies, in fact, even to the weaker requirement that for every two states $\rho_b$ and $\rho_{b'}$ there be a measurement distinguishing them one from another.

2. **Permutability.** The quantum circuit $\mathcal{C}$ might carry out any unitary operation on its quantum state, and in particular an arbitrary permutation on its classical states. Therefore, we need to be able to carry out any permutation of subspaces $A_b$ by a unitary transformation.

It is not hard to demonstrate an efficient encoding that satisfies this permutability condition, without distinguishability: take the subspaces spanned by the basis vectors in the sets $A_b = \{x = (x_1 \cdots x_n) \in \{0, 1\}^n : x_1 = 0 \text{ or } (x_2 \cdots x_n) = b\}$ for $b \in \{0, 1\}^{n-1}$ and let $\rho_b$ be a mixed state over $A_b$.

These subspaces $A_b$ (and the states $\rho_b$) can be permuted by unitary transformations that permute the vectors $|1, x_2, \ldots, x_n\rangle$. However, they are not distinguishable one from another. (Any measurement gives probability distributions that differ only by an exponentially small amount.)

We show that it is not possible to construct an encoding that satisfies both conditions simultaneously. We do this first for classical encodings, then for quantum.

4.2. **CLASSICAL IMPOSSIBILITY RESULT.** To explain the challenge in the argument, let us first describe a simpler, classical analogue of our problem. For simplicity, we only consider the one-clean-qubit case.

We have a classical circuit taking inputs in $\{0, 1\}^n$. This circuit is composed of reversible gates and executes a permutation of $\{0, 1\}^n$. In its starting state, one bit is 0 and the rest are random. Thus, we have a uniform distribution over a set of size $2^{n-1}$ (all inputs with the first bit 0). Executing a reversible gate permutes $\{0, 1\}^n$, mapping this uniform distribution to the uniform distribution over another set of size $2^{n-1}$. The question is: How many sets $A_i \subset \{0, 1\}^n$ of size $2^{n-1}$ can we have so that they are both distinguishable and permutable?

**Theorem 2.** Let $A_1, \ldots, A_M$ be permutable subsets of $\{0, 1\}^n$ of size $2^{n-1}$. If, for some $i, j$, \[\frac{|A_i \cap A_j|}{|A_i|} \leq 1 - \epsilon,\] then $M \leq \frac{2}{\epsilon} n$. 

Thus, if we require that the sets are distinguishable with probability $1/n^c$, we can only have $M = 2^{n^{c+1}}$ sets. This means that the number of bits $m$ that we can simulate is at most $\log M = (c + 1) \log n + 1 = O(\log n)$.

Since the classical case is included only for guidance, we defer its proof to the appendix.

4.3. MAIN THEOREM: QUANTUM IMPOSSIBILITY RESULT. Let $M = 2^m$ be the total number of basis states of the ideal quantum computer being simulated. To simulate the computation, we encode each basis state by an arbitrary subspace $X$ of dimension $2^n - 1$ (or $2^n - k$) within the Hilbert space $\mathbb{C}^{2^n}$ of the computer. (The arbitrary subspaces replace the subsets $A_i$—equivalently axis-parallel subspaces—of the classical case.) The circuit of course can perform not just permutations of the basis, but general unitary operations. In sharp contrast with the classical case, two subspaces of half the dimensionality of the space typically will not intersect. Nevertheless, the large dimension of the subspaces imposes strict constraints on an operator which must permute them; the difficulty is in formulating the incompatibility of these requirements when the number of subspaces is large and the subspaces are required to be very distinct.

If the computer has $k$ clean qubits, then the encodings are subspaces $X$ of dimension $2^n - k$. Let $X$ be the set of these $M$ subspaces $\{X\}$. We show that no oblivious simulation is possible unless $n$ is exponentially larger than $m$ or the number of clean qubits available is at least a constant fraction of $m$:

**THEOREM 3.** If $X$ satisfies the following two constraints:

1. for every permutation $\pi$ on $M$ subspaces, there is a unitary transformation $f_\pi$ such that $f_\pi(X) = \pi(X)$ for all $X \in X$;
2. $\dim(X \cap Y)/\dim(X) < 1 - 1/\text{poly}(m)$ for any $X, Y \in X, X \neq Y$;

then $m \leq (2k + \log n)(1 + o(1))$.

The first constraint is permutability, and the second follows from distinguishability, as shown in:

**LEMMA 1.** Let $X$ and $Y$ be subspaces of dimension $2^{n-k}$ and let $\rho_X$ and $\rho_Y$ be the corresponding density matrices. If $\dim X \cap Y / \dim X = 1 - \epsilon$, then applying any measurement to $\rho_X$ and $\rho_Y$ results in output distributions which differ in variation distance by at most $2\epsilon$.

**PROOF.** Denote $Z = X \cap Y$. Let $X'$ and $Y'$ be the orthogonal complements of $Z$ in $X$ and $Y$; note $\epsilon = (\dim X')/(\dim X) = (\dim Y')/(\dim Y)$. Also,

$$\rho_X = (1 - \epsilon)\rho_Z + \epsilon\rho_{X'}, \quad \rho_Y = (1 - \epsilon)\rho_Z + \epsilon\rho_{Y'}.$$  

For any measurement on a density matrix $(1 - \epsilon)\rho_1 + \epsilon\rho_2$, the output distribution is the convex combination (with coefficients $1 - \epsilon, \epsilon$) of the output distributions for density matrices $\rho_1$ and $\rho_2$. The output distributions for $\rho_X$ and $\rho_Y$ are therefore within variation distance $\epsilon$ of that for $\rho_Z$. \(\square\)

Theorem 3 and Lemma 1 together imply

**COROLLARY 1.** If there exists a distinguishable and permutable encoding of the basis states of an $m$-qubit pure state quantum circuit using $k$-clean qubit states on $n$ qubits, then $m \leq (2k + \log n)(1 + o(1))$.  

This means that an oblivious simulation of $m$-qubit pure state quantum circuits is impossible unless $m \leq (2k + \log n)(1 + o(1))$.

5. Proof Outline

Recall that for every permutation $\pi$ on the $M$ subspaces in $\mathcal{X}$, there is a unitary transformation $f_\pi$ such that $f_\pi(X) = \pi(X)$ for all $X \in \mathcal{X}$.

(I) Suppose that the operators $f_\pi$ form a group representation of $S_M$, that is, $f_{\pi\sigma} = f_\pi f_\sigma$ for all $\pi$ and $\sigma \in S_M$. Assume further that the representation is irreducible.

We would like to prove that $m \leq (2k + \log n)(1 + o(1))$, which is equivalent to $M \leq (n2^{2k})^{1+o(1)}$. For a contradiction, assume that this is not true. Then, we get an upper bound on the dimensionality $2^n$ of the representation, in terms of $M$. This upper bound implies a constraint on the Young diagram of this representation, namely that the first row or column must be very long (almost $M$).

Fixing any $X \in \mathcal{X}$, we must be able to permute all other subspaces in $\mathcal{X}$ arbitrarily. Thus, any $X \in \mathcal{X}$ must be a representation of $S_{M-1}$, the group of permutations that fix $X$. We decompose the representation of $S_M$ into a direct sum of irreducible representations of $S_{M-1}$. One of those representations (the one obtained by deleting an inside corner of the original Young diagram) has dimension that is almost the dimension of the entire representation. This means that the dimension of $X$ is either almost $2^n$ (if it includes the largest representation of $S_{M-1}$) or is much smaller than $2^n$. This contradicts $\dim X = 2^{n-k}$. This forms case 1 of the proof.

In case 2, we handle the situation in which the operators $f_\pi$ still form a group representation of $S_M$, but this representation is reducible. We decompose this representation into irreps. Similarly, we decompose $X, Y \in \mathcal{X}$ into irreps of $S_{M-1}$ with each irrep of $S_{M-1}$ contained in some irrep of $S_M$. Using lemmas from case 1, it follows that each irrep of $S_{M-1}$ is either very big (has a dimension that is either almost the dimension of the corresponding irrep of $S_M$) or very small (has a dimension which is much smaller than the dimension of the corresponding irrep of $S_M$). We then show that $\dim X \cap Y$ is almost the dimension of $X$, by lower-bounding the dimension of the intersection between big irreps contained in $X$ and big irreps contained in $Y$. This contradicts the distinguishability requirement.

(II) Actually the situation is more complicated.

Since $f_\pi f_\sigma(X) = f_\pi(\sigma(X)) = \pi \sigma(X)$, the transformation $f_\pi f_\sigma$ permutes the subspaces in the same way as $f_{\pi\sigma}$. However, this does not mean that $f_\pi f_\sigma = f_{\pi\sigma}$. In other words, the operators $f_\pi$ may not form a representation.

This can arise in several ways. Since every $X$ is a subspace, $-f_\pi f_\sigma(X)$ is the same as $f_\pi f_\sigma(X)$. So, we might have $f_\pi f_\sigma = -f_{\pi\sigma}$ for some $\pi$ and $\sigma$. In general, we can also have more complicated transformations $g$ such that $g(X) = X$ for all $X \in \mathcal{X}$. Then, we might have $f_\pi f_\sigma = g f_{\pi\sigma}$ for some $\pi$ and $\sigma$.

Case 3 of the proof handles the situation when $f_\pi f_\sigma \neq f_{\pi\sigma}$ for some $\pi$ and $\sigma \in S_M$. Let $G$ be the group of transformations that map every subspace $X \in \mathcal{X}$ to itself. Then, $f_\pi f_\sigma f_{\pi\sigma}^{-1}$ is an element of $G$ for any $\pi, \sigma \in S_M$. We would like to modify $f$ so that this element becomes the identity for all $\pi$ and $\sigma \in S_M$. Then, $f_{\pi\sigma} = f_{\pi\sigma}$, that is, $f_{\pi\sigma}$ would form a representation of $S_M$ and we would be able to analyze this representation similarly to the previous section.

If for each $\pi$ we select any element of $G$, call it $g_\pi$, and form the transformation $f_\pi = g_\pi f_\pi$, then these transformations still implement the same permutation $\pi$ of...
What we will do is show how to select the $g_π$ so that on each irrep of $G$, the $f'$ compose as a representation, up to a phase factor: $f'_{πσ} = c_{π,σ,i} f'_{π} f'_{σ}$ for some unit $c_{π,σ,i} ∈ \mathbb{C}$.

The next step is eliminating the phase factors $c_{π,σ,i}$. This is done by lifting these transformations to a space of at most quadratically larger size, in which the phases cancel and the transformations compose as a representation.

Case 3 occupies the bulk of the technical work.

6. Proof of the Main Theorem: Case 1

In this section we consider the case that $f_π$ form a representation ($f_π f_σ = f_πσ$ for all $π$ and $σ$) and this representation is irreducible.

**Lemma 2.** Let $V$ be an irreducible representation of $S_M$ of dimension $N ≤ 2^n$ and $X$ be a subspace fixed by a subgroup isomorphic to $S_{M−1}$. Then, the dimension of $X$ is either at least $(1 − \frac{2^n}{M})N$ or at most $\frac{2^n}{M}N$.

**Proof.** Let $λ$ be the Young diagram of $V$. We show that $λ$ has either a large first row or a large first column.

**Lemma 3.** Either the first row or the first column of $λ$ is of length at least $M − n$.

**Proof.** This follows from a theorem by Rasala [1977]:

**Theorem 4** [RASALA 1977; PP. 151–152]

1. Let $A ≤ M/2$ and $ρ$ be an irreducible representation of $S_M$ such that the first row of the Young diagram of $ρ$ is of length exactly $M − A$. Then,

   $$\dim ρ ≥ ϕ_A(M)$$

   where $ϕ_A(M) = \frac{M−2A+1}{M−A+1} \big(\frac{M}{A}\big)$ is the dimension of the irreducible representation corresponding to the partition $(M − A, A)$.

2. If $ρ$ is an irreducible representation with both the first row and the first column of length at most $M/2$, then

   $$\dim ρ ≥ ϕ_{\lfloor M/2\rfloor}(M).$$

This theorem means that any representation with the first row of the Young diagram having length at most $M − A, A ≤ M/2$ has dimension at least

$$\min_{B:A≤B≤M/2} ϕ_B(M).$$

A simple calculation shows that this expression is minimized by $B = A$ if $A ≤ M/2 − c\sqrt{M}$ for some constant $c$ and $B = \lfloor M/2 \rfloor$ if $A > M/2 − c\sqrt{M}$.

---

1 For case 1, $V$ is the entire state space of $C'$ and $N$ is always equal to $2^n$. We however, state the lemma in a more general version, with $N ≤ 2^n$, because this version will be useful in the proofs for cases 2 and 3.

2 In the second case, the lowest-dimensional representation actually has first row shorter than $M − A$. This is quite surprising because, in most cases, removing a square from the first row of a Young diagram and adding a square somewhere else increases the dimension.
To deduce our lemma, assume that the Young diagram of an irreducible representation of $S_M$ has both first row and column of length at most $M - A$. We show that $N \geq 2^A$. Consider two cases:

Case 1. $A \geq M/2 - c\sqrt{M}$.

Notice that $A \leq M - \sqrt{M}$ because otherwise the Young diagram would fit into a square with a side less than $\sqrt{M}$ and area less than $M$. Theorem 4 implies that

$$\dim \rho \geq \varphi_{\lfloor M/2 \rfloor}(M) = \frac{1}{M} \left( \frac{M}{\lfloor M/2 \rfloor} \right) = \Omega \left( \frac{2^M}{M \sqrt{M}} \right) = 2^{M - \frac{3M}{2} + O(1)} > 2^A.$$ 

Case 2. $A \leq M/2 - c\sqrt{M}$. Then,

$$\dim \rho \geq \varphi_{A}(M) = \frac{M - 2A + 1}{M - A + 1} \left( \frac{M}{A} \right) \geq \frac{1}{M} \left( \frac{M}{A} \right)^A = \frac{1}{M} \left( \frac{M}{2A} \right)^A 2^A \geq 2^A. \quad \square$$

Another lower bound on the longest row or column of a low dimension representation (for a different range of parameters) was given by Mishchenko [1996].

As described in Section 2.4, $V$ decomposes into irreducibles of $S_{M-1}$ in the following way:

$$V = \bigoplus \rho_{\lambda^-}$$

where $\lambda^-$ ranges over all shapes of size $M - 1$ that can be obtained by deleting an "inside corner" from $\lambda$. If $\lambda$ has a long first column or row, one of the $\lambda^-$ has dimension that is almost the same as the dimension of $\lambda$.

**Lemma 4.** Suppose the shape $\lambda$ has a first row (column) of length $|\lambda| - \ell$ for $\ell < |\lambda|/2$. Let $\lambda_1$ denote the "$\lambda^-" obtained by deleting the last element of the first row (column). Then $\dim(\rho_{\lambda_1}) \geq \frac{|\lambda| - 2\ell}{|\lambda|} \dim(\rho_{\lambda})$.

**Proof.** Consider the ratio $\frac{\dim \rho_{\lambda_1}}{\dim \rho_{\lambda}} = \frac{1}{|\lambda|} \prod_{x \in X, h_x} h_x$. In the last ratio, $h_x$ for points $x$ outside of the first row (column) are identical in the numerator and denominator. Moreover for each $x$ in the first row (column) in the numerator other than the very last point (which contributes a factor of 1 in the numerator and is absent in the denominator), the ratio between its contributions in the numerator and denominator is $\frac{h_x}{h_x - 1}$ where $h_x$ is the length of its hook in $\lambda$.

Since there are only $\ell$ squares below the first row, at most $\ell$ squares in the first row have another square below them. Thus, none of $(|\lambda| - \ell) - \ell = |\lambda| - 2\ell$ rightmost squares in the first row of $\lambda$ has another square below it. The rightmost of them is absent in $\lambda_1$, the others have hook-lengths equal to $2, 3, \ldots, |\lambda| - 2\ell$ in $\lambda$ and $1, 2, \ldots, |\lambda| - 2\ell - 1$ in $\lambda_1$. This lower-bounds the ratio by $\prod_{i=1}^{|\lambda| - 2\ell} \frac{i + 1}{i} = |\lambda| - 2\ell$. Therefore, $\frac{\dim \rho_{\lambda_1}}{\dim \rho_{\lambda}} \geq \frac{|\lambda| - 2\ell}{|\lambda|}$. \quad \square

Let $W = \rho_{\lambda_1}$. Since $X$ and $W$ are invariant under $S_{M-1}$, $W \cap X$ must be invariant as well. Since $W$ is irreducible, $W \cap X$ must be either $0$ or $W$. If $W \cap X = W$, then $W \subseteq X$, $\dim X \geq \dim W \geq \dim W \geq (1 - \frac{2\ell}{|\lambda|})N$ by Lemmas 4 and 3. If $W \cap X = 0$,
then \( \dim X + \dim W \leq N \) and Lemmas 4 and 3 imply \( \dim X \leq N - \dim W \leq \frac{2n}{M} N \). This completes the proof of Lemma 2. \( \square \)

To deduce Theorem 3 from Lemma 2, let \( X \in \mathcal{X} \). Then, \( X \) is fixed by a subgroup isomorphic to \( S_{M-1} \) and \( \dim X = \frac{N}{2} \). By Lemma 2, \( \dim X \leq \frac{2n}{M} N \). Therefore, \( \frac{N}{2} \leq \frac{2n}{M} N \), implying \( M \leq 2^{k+1} n \). Taking logarithms of both sides gives \( m = \log M \leq \log n + k + 1 \). This proves Theorem 3 in case 1.

7. Proof of the Main Theorem: Case 2

In this section, we consider the case that \( f_\pi \) form a representation \( (f_\pi f_\sigma = f_\pi \sigma \) for all \( \pi \) and \( \sigma \) but this representation is not irreducible.

In case 1, it was not possible to have a subspace \( X \) of dimension \( \frac{\dim V}{2} \) that is fixed by \( S_{M-1} \). If we do not require the representation to be irreducible, there are many ways to construct such a subspace \( X \) (and a full set \( \mathcal{X} \) of \( M \) permutable subspaces). However, in any such construction, any two of the subspaces must have a large-dimensional intersection.

For the proof of the theorem in case 2, first suppose that \( n \geq M/4 \). In this case, taking logarithms of both sides we find that \( m \leq \log n + 2 \) and hence \( m \leq (k + \log n)(1 + o(1)) \) as desired.

For the rest of this section suppose that \( n < M/4 \).

**Lemma 5.** Let \( N < 2^n \). Let \( f_\pi \) be an \( N \)-dimensional representation of \( S_M \) that acts as a permutation representation on a collection of \( M \) subspaces \( \mathcal{X} \). Then, for any \( X, Y \in \mathcal{X} \),

\[
\dim X - \dim X \cap Y \leq \frac{4n}{M} N.
\]

**Proof.** We decompose \( V \) into irreducible representations of \( S_M \):

\[
V = \bigoplus_{i=1}^{\ell} W_i \otimes E_i
\]

where the \( W_i \) are nonisomorphic, and each \( E_i \) is a space on which \( S_M \) acts as the identity, with \( \dim E_i \) equal to the number of times \( W_i \) is contained in \( V \).

We also decompose \( V \) into irreducibles of \( S_{M-1}^X \) (the subgroup of \( S_M \) that fixes \( X \)) in a similar way: \( V = \bigoplus_{i=1}^{\ell'} U_i \otimes F_i \) where \( U_i \) are distinct irreducibles of \( S_{M-1}^X \) and each \( F_i \) is a space on which \( S_M \) acts as the identity.

**Lemma 6.** For each \( U_i \), there is at most one \( j \) such that \( U_i \) is the largest irreducible of \( S_{M-1}^X \) contained in \( W_j \).

**Proof.** By Lemma 3, the first row or the first column of any \( W_j \) is of the length at least \( M - n \geq \frac{3}{2} M \). Consider the irreducible \( \rho_j \) of \( S_{M-1}^X \) with the Young diagram obtained by deleting the last square from the Young diagram of \( W_j \). By Lemma 4,

\[
\dim \rho \geq \frac{|\lambda| - 2l}{|\lambda|} \dim W_j \geq \frac{M - 2n}{M} \dim W_j > \frac{1}{2} \dim W_j.
\]

Therefore, \( \rho_j \) is the largest of the irreducibles of \( S_{M-1}^X \) contained in \( W_j \). If \( W_j \) and \( W_{j'} \) are different irreducibles of \( S_M \), they must have different Young diagrams \( \lambda_j \) and \( \lambda_{j'} \). Then, the diagrams obtained by deleting the last square of the
We reorder $U_i$ so that, for each $i \in \{1, \ldots, p\}$, $U_i$ is the largest irreducible of $W_i$. If there are $S_{M-1}^k$-irreducibles $U_i$ that are not equal to the largest irreducible of any $W_i$ ($i \in \{1, \ldots, p\}$), we denote them $U_{p+1}, U_{p+2}, \ldots$.

Consider the intersection of $U_i \otimes F_i$ and $W_i \otimes E_i$. This intersection must be of the form $U_i \otimes K_i$ for some $K_i$. Let $L_i$ be the complement of $K_i$ in $F_i$. (If $i > p$, then $K_i = \{\overrightarrow{0}\}$ and $L_i = F_i$.) Define

$$X' = \sum_{i=1}^{p} U_i \otimes K_i,$$

$$X'' = \sum_{i=1}^{r} U_i \otimes L_i.$$  

Then $X = X' \oplus X''$. Let $\pi \in S_M$ be such that $\pi(X) = Y$. Define $Y' = f_\pi(X')$, $Y'' = f_\pi(X'')$. Then, $Y = Y' \oplus Y''$. The Lemma 5 follows from two following lemmas:

**LEMMA 7**

$$\dim(X') - \dim(X' \cap Y') \leq \frac{2n}{M}N.$$  

**PROOF.** We have $Y' = \bigoplus_{i=1}^{p} f_\pi(U_i) \otimes K_i$. By Lemmas 3 and 4, the dimension of $U_i$ (and $f_\pi(U_i)$) is at least $\frac{M-2n}{M} \dim(W_i)$. Therefore, $\dim(U_i) - \dim(U_i \cap f_\pi(U_i)) \leq \frac{2n}{M} \dim(W_i)$. The lemma follows by summation over all $i$.  

**LEMMA 8**

$$\dim(X'') \leq \frac{2n}{M}N.$$  

**PROOF.** First, notice that $U_i \otimes L_i$ is orthogonal to $\sum_{j=1}^{p} U_j \otimes E_j$. (It is orthogonal to $U_i \otimes E_j$ because $(U_i \otimes E_j) \cap (U_i \otimes F_i) = U_i \otimes K_i$ and $L_i$ is the complement of $K_i$ in $F_i$. It is orthogonal to $U_j \otimes E_j$ for $j \neq i$ because $U_i$ and $U_j$ are different irreducibles of $S_{M-1}$ (Lemma 6).

Therefore, $U_i \otimes L_i$ is contained in $\sum_{j=1}^{p} (W_j - U_j) \otimes E_j$ and $X''$ (which is the direct sum of $U_i \otimes L_i$ over all $i$) is contained in $\sum_{j=1}^{p} (W_j - U_j) \otimes E_j$ as well. We have

$$\dim X'' \leq \sum_{j=1}^{p} (\dim W_j - \dim U_j) \dim E_j \leq \frac{2n}{M}N,$$

with the last inequality following from Lemmas 3 and 4.  

If $f_\pi$ form a representation, Lemma 5 almost immediately implies Theorem 3. Assuming $n < M/4$, we have

$$\frac{\dim X \cap Y}{\dim X} = \frac{\dim X - (\dim X - \dim X \cap Y)}{\dim X} \geq \frac{2^{n-k} - \frac{4n}{M} 2^n}{2^{n-k}} = 1 - \frac{2^{k+2n}}{M}.$$  

If this is at most $1 - \frac{1}{\text{poly}(m)}$, then $m \leq (k + \log n)(1 + o(1))$.  


8. Proof of the Main Theorem: Case 3

In this section, we consider the case that \( f_\pi \) do not form a representation (\( f_\pi f_\sigma \neq f_{\pi\sigma} \) for some \( \pi \) and \( \sigma \)).

8.1. PROOF OUTLINE. Let \( G \) be the group of transformations that map every subspace \( X \in \mathcal{X} \) to itself. Then, \( f_\pi f_\sigma f_{\pi^{-1}} \) is an element of \( G \) for any \( \pi, \sigma \in S_M \). We would like to modify \( f \) so that this element becomes the identity for all \( \pi \) and \( \sigma \in S_M \). Then, \( f_\pi f_\sigma = f_{\pi\sigma} \), i.e., \( f_\pi \) would form a representation of \( S_M \) and we would be able to analyze this representation similarly to the previous section.

To achieve this, we look at \( \mathbb{C}^{2^n} \) as a representation of \( G \) and express \( \mathbb{C}^{2^n} \) as \( V_1 \oplus V_2 \cdots \oplus V_k \), with \( V_i \) corresponding to different types of irreducible representations of \( G \).

Then, we compose each \( U_\pi \) with an appropriate \( g_\pi \in G \). The resulting transformation \( f'_\pi = g_\pi f_\pi \) still implements the same permutation \( \pi \) of \( \mathcal{X} \) because \( g_\pi \) maps every \( X \in \mathcal{X} \) to itself. We can choose the transformations \( g_\pi \) so that, on every \( V_i \), \( f'_\pi f_\pi \) is the same as \( f'_\pi \) up to a phase (\( f'_\pi = c_{\pi,\sigma,i} f'_\sigma f_\sigma \) for some unit \( c_{\pi,\sigma,i} \in \mathbb{C} \)).

The next step is eliminating the phase factors \( c_{\pi,\sigma,i} \). This is done by considering a larger space \( V_1 \otimes V_1^* + \cdots + V_k \otimes V_k^* \) and transformations \( f''_\pi = f'_\pi \otimes (f'_\pi)^* \) on this larger space. Then, the phase factors \( c_{\pi,\sigma,i} \) (from \( f' \)) and \( c_{\pi,\sigma,i}^* \) (from \( (f')^* \)) cancel out and we get \( f''_\pi = c_{\pi,\sigma,i} c_{\pi,\sigma,i}^* f''_\sigma f'^*_\sigma = f''_\sigma f'^*_\sigma \). Thus, \( f''_\pi \) form a representation of \( S_M \) on the linear space \( V_1 \otimes V_1^* + \cdots + V_k \otimes V_k^* \). This representation can be then analyzed similarly to case 2.

8.2. REPRESENTATION UP TO PHASES \( c_{\pi,\sigma,i} \). Let \( G \) be the group of unitary transformations that fix every one of the subspaces \( X \in \mathcal{X} \).

LEMMA 9. \( G \) is a compact Lie group.

PROOF. \( G \) is clearly a Lie group. To show that \( G \) is compact, we first observe that \( G \) is a subgroup of \( SU(2^n) \), the group of all unitary transformations on a \( 2^n \)-dimensional space and \( SU(2^n) \) is compact. Therefore, it suffices to show that \( G \) is a closed subset of \( SU(2^n) \).

Let \( U_1 \in G, U_2 \in G, \ldots \) and assume that \( U = \lim_{n \to \infty} U_n \) exists. We have to show that \( U \in G \) which is equivalent to \( U(X) = X \) for every \( X \in \mathcal{X} \). If \( U_1 \in G, U_2 \in G, \ldots \), then \( U_1(X) = X, U_2(X) = X, \ldots \). Therefore, \( U(X) = X \).

Therefore, the representations of \( G \) are completely reducible and Schur’s lemma applies to irreducible representations. \( \mathbb{C}^{2^n} \) is a representation of \( G \) and all \( X \in \mathcal{X} \) are invariant subspaces. (They are fixed by every element of \( G \) according to the definition of \( G \).) These invariant subspaces decompose into irreducible invariant subspaces.

We decompose \( \mathbb{C}^{2^n} \) into irreducible invariant subspaces and split those irreducible invariant subspaces into equivalence classes consisting of isomorphic irreducible subspaces. Let \( E_1, \ldots, E_k \) be these equivalence classes. Let \( V_i \) be the subspace of \( \mathbb{C}^{2^n} \) spanned by all the irreducible subspaces in \( E_1 \) (i.e., the subspace spanned by all the vectors belonging to at least one subspace in \( E_1 \)). Let \( V_2, \ldots, V_k \) be defined similarly.
CLAIM 1. If $i, j \in \{1, \ldots, k\}$ and $i \neq j$, then $V_i \perp V_j$.

Therefore, $C^2 = V_1 \oplus V_2 \oplus \cdots \oplus V_k$. Next, we show that transformations $f_\pi$ map each $V_i$ to some (possibly different) $V_i'$.

CLAIM 2. Let $V$ be a $G$-invariant subspace. Then, $f_\pi(V)$ is $G$-invariant as well. If $V$ is irreducible, $f_\pi(V)$ is irreducible. Moreover, if $V$ and $V'$ are two $G$-isomorphic irreducible subspaces, $f_\pi(V)$ and $f_\pi(V')$ are isomorphic as well.

PROOF. Let $g \in G$. The map $g \rightarrow f_\pi g f_\pi^{-1}$ is an automorphism of $G$. If $V$ is invariant under the action of $g$, $f_\pi(V)$ is invariant under the action of $f_\pi g f_\pi^{-1}$. Therefore, if $V$ is $G$-invariant, so is $f_\pi(V)$.

If $f_\pi(V)$ is not irreducible, it decomposes into two or more $G$-invariant subspaces: $f_\pi(V) = W_1 \oplus W_2$. Then, $f_\pi^{-1}(W_1)$ is $G$-invariant as well, implying that $V$ is not irreducible.

Finally, let $h : V \rightarrow V'$ be a $G$-isomorphism of $V$ and $V'$ (an isomorphism that commutes with the action of $G$). Let $h' : f_\pi(V) \rightarrow f_\pi(V')$ be defined by $h' = f_\pi h f_\pi^{-1}$. Then, for any $g = f_\pi g f_\pi^{-1}$, we have

$$h' = f_\pi g f_\pi^{-1}$$

and every $g \in G$ can be expressed in the form $f_\pi g' f_\pi^{-1}$. Therefore, $h'$ is a $G$-isomorphism of $f_\pi(V)$ and $f_\pi(V')$.

Remark 1. $V$ does not have to be isomorphic to $f_\pi(V)$ as a representation of $G$. $f_\pi$ establishes the isomorphism of $g$ on $V$ with $f_\pi g f_\pi^{-1}$ on $f_\pi(V)$, but $f_\pi g f_\pi^{-1}$ is not necessarily equal to $g$ on $f_\pi(V)$.

CLAIM 3. For every $i \in \{1, \ldots, k\}$ there is an $i'$ such that $f_\pi(V_i) = V_i'$.

PROOF. By Claim 2, every two isomorphic irreducible subspaces get mapped to isomorphic irreducible subspaces. Therefore, all subspaces in $E_i$ get mapped to subspaces in the same $E_i'$ and $f_\pi(V_i) \subseteq V_i'$. Similar reasoning applied to $f_\pi^{-1}$ implies $f_\pi^{-1}(V_i') \subseteq V_i$.

For each $i \in \{1, \ldots, k\}$, $V_i$ is the direct sum of some number of isomorphic irreducible subspaces: $V_i = V_{i1} \oplus V_{i2} \oplus \cdots \oplus V_{ik}$. We fix $G$-isomorphisms $h_{ijj'}$ between $V_{ij}$ and $V_{ij'}$ so that $h_{ijj'}h_{ijj'}' = h_{ijj''}$.

LEMMA 10 [FULTON AND HARRIS 1991; SCHUR’S LEMMA]. Let $V$ and $V'$ be an isomorphic irreducible representations of $G$. Then, for any two $G$-isomorphisms $h : V \rightarrow V'$ and $h' : V \rightarrow V'$, there is a constant $c$ such that $h = ch'$.

Therefore, each $h_{ijj'}$ is unique up to a multiplicative constant. The isomorphisms can be made to compose properly by adjusting these constants. Note that $h_{ijj}$ is the identity.

CLAIM 4. $W \subseteq V_i$ is an irreducible invariant subspace (under the action of $G$) if and only if

$$W = \{a_j x + a_{j+1} h_{ij(j+1)}(x) + \cdots + a_k h_{ijk}(x) | x \in V_{ij}\}$$

for some $j \in \{1, \ldots, j_i\}$ and $a_j, \ldots, a_k \in \mathbb{C}$. 

A. AMBAINS ET AL.
Proof. “If” part:
Invariance: Let \( g \in G \).
\[
g \left( \sum_{\ell=j}^{j'} a_{\ell} h_{ij\ell}(x) \right) = \sum_{\ell=j}^{j'} a_{\ell} g(h_{ij\ell}(x)) = \sum_{\ell=j}^{j'} a_{\ell} h_{ij\ell}(g(x))
\]
because each \( h_{ij\ell} \) is a \( G \)-isomorphism.

\( W \) is irreducible because, if \( W_1 \subseteq W \) and \( W_1 \) is \( G \)-invariant, then
\[
\{ x | a_j x + a_{j+1} h_{ij(j+1)}(x) + \cdots + a_{j'} h_{ijj}(x) \in W_1 \}
\]
is a \( G \)-invariant subspace of \( V_{ij} \); but \( V_{ij} \) is irreducible and \( \dim(W) = \dim(V_{ij}) \).

“Only if” part:
Let \( W \) be an irreducible invariant subspace of \( V_i \). Let \( x' \in W \). Then, we can write \( x' = x_j' + \cdots + x_{j'}' \), \( x_j' \in V_{i1}, \ldots, x_{j'}' \in V_{ij} \). If \( x_j' \neq x'' \in W \), then for any index \( j \), \( x_j' \neq x_j'' \) or \( x_j' = x_j'' = 0 \). (Otherwise, \( W \cap \bigoplus_{\ell \neq j} V_{i\ell} \) is a nontrivial subspace of \( W \). It is \( G \)-invariant because \( V_i \) and \( V_{ij} \) are \( G \)-invariant. Contradiction with the irreducibility of \( W \).)

Let \( j \) be the smallest index for which there is an \( x_j' \in W \) with \( x_j' \neq 0 \). Then, for every \( x \in V_{ij} \), there is an \( x' \in W \) with \( x_j' = x \). (For, if \( A \) and \( B \) are invariant subspaces of a unitary representation, the projection of \( A \) onto \( B \) is invariant. Apply this with \( A = W \) and \( B = V_{ij} \), then use the irreducibility of \( V_{ij} \).)

The above considerations allow us to define the mapping \( h_{jj'} : V_{ij} \to V_{ij'} \) by \( h_{jj'}(x_j') = x_j' \). By the definition, \( h_{jj'}(g(x_j')) = h_{jj'}((g(x_j'))_j) = (g(x_j'))_{j'} = g(x_j') = g(h_{jj'}(x_j')) \), so \( h_{jj'} \) is a \( G \)-isomorphism. By Schur’s lemma, this implies that \( h_{jj'} = a_{jj'} h_{jj'} \) for some \( a_{jj'} \in \mathbb{C} \).

In general, if \( W \) is any irreducible \( G \)-invariant subspace of \( \mathbb{C}^{2^n} \), then \( W \) must be in the form described by Claim 4 for some \( i \). (\( W \) belongs to some equivalence class \( E_i \) and therefore is contained in the corresponding \( V_i \).)

For each \( V_{i1} (i \in \{1, \ldots, k\}) \), we fix an orthonormal basis \( v_{i1}, \ldots, v_{it} \). This also fixes a related basis \( h_{i11}(v_{i1}), \ldots, h_{i11}(v_{it}) \) for each \( V_{ij} \). Moreover, we also get a similar basis
\[
\sum_{\ell=j}^{j'} a_{\ell} h_{i1\ell}(v_{i1}), \ldots, \sum_{\ell=j}^{j'} a_{\ell} h_{i1\ell}(v_{it})
\]
for every \( G \)-invariant irreducible \( W \subseteq V_i \) because any such \( W \) can be written in the form given by Claim 4. We call these bases designated.

This designated basis is exactly the basis for \( W \) that can be obtained by applying the isomorphism between \( V_{i1} \) and \( W \) to the basis for \( V_{i1} \). Moreover, if \( W, W' \) are two isomorphic irreducible subspaces, the designated basis for \( W \) is mapped to the designated basis for \( W' \) by the isomorphism between \( W \) and \( W' \).

We are going to impose the following condition on \( f'_{\pi} \):

Condition. Let \( W \) be an irreducible representation of \( G \) and \( w_1, \ldots, w_l \) be the designated basis of \( W \). Let \( w'_1, \ldots, w'_l \) be the designated basis of \( f'_{\pi}(W) \). Then, there exists \( c \in \mathbb{C}, |c| = 1 \) such that \( f'_{\pi}(w_1) = cw'_1, \ldots, f'_{\pi}(w_l) = cw'_l \).

Next, we show that this condition suffices to guarantee \( f_{\pi} f_{\pi}' = c_{\pi,\sigma} f_{\pi,\sigma}' \) on every \( V_i \) and that any \( f_{\pi} \) that permutes \( X \in \mathcal{X} \) without satisfying this condition can
be transformed into $f'_v$ that satisfies the condition and still permutes the subspaces in the same way.

First, we show that it is enough to ensure that the designated basis of $V_{i1}$ is mapped correctly for every $i \in \{1, \ldots, k\}$.

**Claim 5.** Assume that the condition is true for $W = V_{i1}$. Then, it is also true for any irreducible $W \subseteq V_i$.

**Proof.** Let $h$ be the isomorphism between $V_{i1}$, $W$. Note that $h$ maps the designated basis of $V_{i1}$ to the designated basis of $W$.

Then (by Claim 2) $f'_v h(f''_v)^{-1}$ is an isomorphism between $f'_v(V_{i1})$ and $f''_v(W)$. We know that there is an isomorphism between these two irreducibles that maps the designated basis of one of them to the designated basis of the other. By Schur’s lemma, any two isomorphisms of irreducible subspaces can differ only by a multiplicative constant $c$. The unitarity of $h(f''_v)^{-1}$ implies that $|c| = 1$.

Therefore, $f'_v h(f''_v)^{-1}$ maps the designated basis of $f'_v(V_{i1})$ to $c$ times the designated basis of $f''_v h(f''_v)^{-1}(f'_v(V_{i1})) = f''_v h(V_{i1}) = f''_v(W)$. We know that $(f''_v)^{-1}$ maps the designated basis of $f''_v(V_{i1})$ to the designated basis of $V_{i1}$ and that $h$ maps the designated basis of $V_{i1}$ to the designated basis of $W$. This implies that $f'_v$ maps the designated basis of $W$ to $c$ times the designated basis of $f''_v(W)$. □

Next, we show how to transform $f_\pi$ into $f'_\pi$ that performs the same permutation $\pi$ of $\mathcal{X}$ and maps the designated basis of every $V_{i1}$ as required.

Let $W_1, \ldots, W_k$ be $f_\pi(V_{i1})$, $\ldots, f_\pi(V_{k1})$. Each of $W_i$ lies within one of $V_1, \ldots, V_k$. Denote this subspace $V'_{i1}$. Then, for $i \neq j$, $V'_i \neq V'_j$. For each $i \in \{1, \ldots, k\}$, we define a unitary transformation $g_{\pi,i}$ on $V_i$ such that $g_{\pi,i} f_\pi$ maps the designated basis of $V_{i1}$ to the designated basis of $f_\pi(V_{i1})$.

By Claim 4, the irreducible subspace $W_i = f_\pi(V_{i1})$ is just

$$\{a_{i,j}x + a_{i,(j+1)}h_{i,j(j+1)}(x) + \cdots + a_{i,j}h_{i,jj}(x)|x \in V_{i1}\}$$

for some $j$. Moreover, the mapping that maps each $v \in W_i$ to its $V'_{i1}$-component is an isomorphism of $W_i$ and $V'_{i1}$ with respect to $G$ (similarly to proof of Claim 4).

Let $v_1, \ldots, v_l$ be the designated basis of $V_{i1}$, $v'_1, \ldots, v'_l$ be $f_\pi(v_1), \ldots, f_\pi(v_l)$ and $v''_1, \ldots, v''_l$ be the $V'_{i1}$ components of $v'_1, \ldots, v'_l$.

Let $w_1, \ldots, w_l$ be the designated basis of $V'_{i1}$ and $g_{\pi,i,j}$ be the unitary transformation on $V'_{i1}$ that maps $v''_i$, $\ldots, v''_l$ to $w_1, \ldots, w_l$. We define a unitary transformation $g_{\pi,i,j}$ (for every $j' \neq j$ on $V'_{i1}$) to be $h_{i,j(j+1)}g_{\pi,i,j} h_{i,jj}$'s. Finally, we take the transformation $g_{\pi,i,j}$ of $V'_{i1}$ that is equal to $g_{\pi,i,j}$ on each $V'_{i1}$. Then, $g_{\pi,i,j}$ maps

$$v'_{i1} = a_{i,j}v''_1 + a_{i,(j+1)}h_{i,j(j+1)}(v''_1) + \cdots + a_{i,j}h_{i,jj}(v''_1)$$

to

$$a_{i,j}g_{\pi,i,j}(v''_1) + a_{i,(j+1)}g_{\pi,i,(j+1)}h_{i,j(j+1)}(v''_1) + \cdots = a_{i,j}g_{\pi,i,j}(v'_1) + a_{i,(j+1)}h_{i,j(j+1)}g_{\pi,i,j}(v'_1) + \cdots = a_{i,j}w_1 + a_{i,(j+1)}h_{i,j(j+1)}(w_1) + \cdots$$

which is exactly the first vector of the designated basis for $W_i$. The same is true for $v''_2, \ldots, v''_l$, implying that $g_{\pi,i,j} f_\pi$ maps the designated basis of $V_{i1}$ to the designated basis of $W_i$.

Now, we take $g_\pi$ that is equal to $g_{\pi,i}$ on each $V_i$ and take $f'_\pi = g_\pi f_\pi$. 

---

526 A. AMBAINIS ET AL.
CLAIM 6. \( g_\pi \) preserves all \( X \in \hat{X} \).

PROOF. By definition, the restriction \( g_\pi |_{V_\pi} \) is equal to \( g_{\pi,i} \), and \( g_{\pi,i} \) clearly preserves \( V_{i_1}, \ldots, V_{i_j} \). Moreover, \( g_{\pi,i} \) (and, hence, \( g_\pi \)) preserves any irreducible subspace \( W \subseteq V_i \) because any such subspace is in the form of Claim 4.

Every \( X \in \hat{X} \) is invariant under \( G \). Therefore, it decomposes into a direct sum of irreducible subspaces. Each of these subspaces is in one of the classes \( E_1, \ldots, E_k \) and, therefore, lies in one of \( V_1, \ldots, V_k \). This means that it is preserved by \( g_\pi \). Therefore, \( X \) which is a direct sum of such irreducible subspaces is preserved by \( g_\pi \) as well. \( \square \)

Hence, \( f'_\pi = g_\pi f_\pi \) realizes the same permutation \( \pi \) of \( X \in \hat{X} \) as \( f_\pi \).

CLAIM 7. On every \( V_i \), \( f'_\pi f'_\sigma = c_{\pi,\sigma,i} f'_{\pi\sigma} \) for some \( c_{\pi,\sigma,i} \in \mathbb{C} \).

PROOF. This is equivalent to showing that \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \) is equal to \( c_{\pi,\sigma,i} \) times the identity. To show that, notice that \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \) maps every subspace \( X \in \hat{X} \) to itself because \((f'_{\pi\sigma})^{-1} \) performs the inverse of the permutation \( \pi \sigma \) on \( \hat{X} \). Therefore, \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \in G \). This means that \( V_{ij} \) are all preserved by \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \).

Moreover, \( f'_\pi \), \( f'_\sigma \) and \( f'_{\pi\sigma} \) all map the designated bases to \( c \)-times designated bases (Claim 5). Therefore, \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \) maps the designated basis of \( V_{ij} \) to \( c \) times the designated basis of \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma(V_{ij}) = V_{ij} \).

It remains to show that \( c \) is the same for all irreducible subspaces \( V_{ij} \) contained in \( V_i \). Let \( c_j \) and \( c_{j'} \) be the values of \( c \) for \( V_{ij} \) and \( V_{ij'} \). Consider the subspace

\[
W = \{ x + h_{ijj'}(x) | x \in V_{ij} \}.
\]

By Claim 4, this is an irreducible \( G \)-invariant subspace. Now, \((f'_{\pi\sigma})^{-1} f'_\pi f'_\sigma \) maps it to

\[
W' = \{ c_j x + c_{j'} h_{ijj'}(x) | x \in V_{ij} \} \]

\[
= \left\{ x + \frac{c_{j'}}{c_j} h_{ijj'}(x) | x \in V_{ij} \right\}.
\]

The \( G \)-invariance of \( W \) means that \( W' = W \) and \( c_j = c_{j'} \).

Therefore, \( c_j \) are all equal. This means that \((f'_{\pi\sigma})^{-1}(x) = c_j x \) for all \( x \in V_i \) because the designated bases of \( V_{ij} \) together form a basis for entire subspace \( V_i \). \( \square \)

Unfortunately, arguments of this type (composing \( f_\pi \) with an appropriate transformation that fixes all \( V_i \) cannot be used to eliminate phases \( c_{\pi,\sigma,i} \).

The reason for this is that there exist so-called projective representations. A projective representation is a set of maps \( f_\pi \) such that \( f_\pi f_\sigma = c_{\pi,\sigma} f_{\pi\sigma} \), \( c_{\pi,\sigma} \in \mathbb{C} \). It is known that the symmetric group has projective representations which are not equivalent to any of the usual representations [Hoffman and Humphreys 1992].

One possible solution would be to use the standard forms of projective representations which are quite well studied [Hoffman and Humphreys 1992]. However, to be able to use them, we would need to show that the multiplicative constants \( c_{\pi,\sigma,i} \) are the same for all \( V_i \) (or show that we can split all \( V_i \) in several groups so that \( c_{\pi,\sigma,i} \) is the same within one group) and we do not know if this is possible.
Our solution is to replace \( f'_\sigma \) by transformations \( f''_\sigma \) on a larger space \( V_1 \otimes V_1^* + \cdots V_k \otimes V_k^* \) so that \( f''_\sigma(f''_\sigma)^* = f''_\sigma f''_\sigma^* \). Then, \( f''_\sigma \) form a representation in the usual sense and we can analyze them similarly to Section 7.

8.3. Fixing the Phase Problem. We split \( V_1, \ldots, V_k \) into equivalence classes \( V_1, \ldots, V_l \). \( V_i \) and \( V_j \) are in one class if there is a \( \pi \in S_M \) such that \( f_\pi(V_i) = V_j \). Let \( W_i \) be the union of all \( V_j \) that belong to \( V_i \). Then, \( f'\pi(W_i) = W_i \) for any \( \pi \in S_M \) (because \( f_\pi \) maps every \( V_j \in V_i \) to some \( V_j' \in V_i \)). Therefore, we can look at each \( W_i \) separately.

Each subspace \( X \) is invariant under any transformation in \( G \) (because we defined \( G \) as the group of transformations that fix all \( X \)). Therefore, \( X \) is a direct sum of irreducible representations of \( G \). Each of those irreducibles is contained in some \( V_i \). Therefore,

\[
X = \bigoplus_{i=1}^k X \cap V_i.
\]

If we group together all \( V_i \) that belong to the same equivalence class \( V_i \), we get

\[
X = \bigoplus_{j=1}^l X \cap W_j, \quad X \cap W_j = \bigoplus_{i \in V_j} X \cap V_i.
\]

The intersections \( X \cap Y \) decompose similarly. So, we can show the following lemma and then obtain Theorem 3 by summation over \( W_j \).

**Lemma 11.** Let \( X, Y \in \hat{X} \). Then, for any \( j \),

\[
\dim X \cap W_j - \dim X \cap Y \cap W_j \leq \sqrt{\frac{4n}{M}} \dim W_j.
\]

**Proof.** To simplify the notation, assume that \( W_j = V_1 \oplus V_2 \oplus \cdots \oplus V_l \).

Consider the linear space \( W'_j = V_1 \otimes V_1^* + \cdots V_l \otimes V_l^* \) and the linear transformations \( f''_\sigma = f''_\sigma \otimes (f''_\sigma)^* \). These linear transformations form a representation because

\[
f''_\sigma \otimes (f''_\sigma)^* = c_{\pi,\sigma,i} f''_\sigma \otimes c_{\pi,\sigma,i}^* (f''_\sigma)^* (f''_\sigma)^* = f''_\sigma \otimes (f''_\sigma)^* (f''_\sigma)^* = f''_\sigma f''_\sigma^*
\]

on every \( V_i \otimes V_i^* \). The counterpart of \( X \) is

\[
X' = \bigoplus_{i=1}^l (X \cap V_i) \otimes (X \cap V_i)^*.
\]

We have \( f''_\sigma(X') = Y' \). (To see this, consider one of \( (X \cap V_i) \otimes (X \cap V_i)^* \). Assume that \( f''_\sigma \) maps \( V_i \) to \( V_i' \). Then, \( f''_\sigma(X) = Y \) implies \( f''_\sigma(X \cap V_i) = Y \cap V_i' \) and

\[
f''_\sigma((X \cap V_i) \otimes (X \cap V_i)^*) = (Y \cap V_i') \otimes (Y \cap V_i')^*.
\]

Combining these equalities for all \( V_i \) gives \( f''_\sigma(X') = Y' \).

In particular, \( f''_\pi(X') = Y' \) means that \( X' \) is invariant under all \( \pi \in S_M \) satisfying \( \pi(X) = X \). Therefore, by Lemma 5,

\[
\dim X' - \dim X' \cap Y' \leq \frac{4n}{M} \dim W_j.
\]
We use this inequality to derive a bound on $\dim X \cap W_j - \dim X \cap Y \cap W_j$. To do this, we relate the dimensions of $X \cap V_i$ and $X' \cap (V_i \otimes V_i^*)$. Remember that
\begin{equation}
X \cap W_j = \bigoplus_{i=1}^l (X \cap V_i),
\end{equation}
and
\begin{equation}
X \cap Y \cap W_j = \bigoplus_{i=1}^l (X \cap Y \cap V_i).
\end{equation}
Let $d_i$ and $d'_i$ be the dimensions of $X \cap V_i$ and $X' \cap (V_i \otimes V_i^*)$. Then, (4) and (5) imply that $\dim X \cap W_j = \sum_{i=1}^l d_i$, $\dim X \cap Y \cap W_j = \sum_{i=1}^l d'_i$, and
\begin{equation}
\dim X \cap W_j - \dim X \cap Y \cap W_j = \sum_{i=1}^l (d_i - d'_i).
\end{equation}
If we look at $V_i \otimes V_i^*$, then $X' \cap (V_i \otimes V_i^*) = (X \cap V_i) \otimes (X \cap V_i)^*$.
This implies $\dim X' \cap (V_i \otimes V_i^*) = d_i^2$ and $\dim X' = \sum_i d_i^2$. Similarly, $\dim X' \cap Y' = \sum_{i=1}^l d_i'^2$.
Let $d$ be the dimension of $V_1$. Then, the dimensions of $V_2, \ldots, V_l$ are $d$ as well because, for every $i \in \{2, \ldots, l\}$, there is a unitary $f_{\pi}$ such that $f_{\pi}(V_1) = V_i$.
Therefore, $\dim W_i = ld$. Also, $\dim W'_i = ld^2$ because $\dim V_i \otimes V_i^* = d^2$ for every $i \in \{1, \ldots, l\}$. Hence, we have
\begin{equation}
\frac{\dim X \cap W_j - \dim X \cap Y \cap W_j}{\dim W_j} = \frac{\sum_{i=1}^l (d_i - d'_i)}{ld} \leq \frac{1}{l} \sum_{i=1}^l \sqrt{\frac{(d_i - d'_i)^2}{d^2}} \leq \frac{1}{l} \sum_{i=1}^l \sqrt{\frac{(d_i - d'_i)(d_i + d'_i)}{d^2}} \leq \frac{1}{l} \sum_{i=1}^l \sqrt{\frac{d_i'^2 - d_i^2}{d^2}}.
\end{equation}
The convexity of the square root implies that this is at most
\begin{equation}
\sqrt{\frac{\sum_{i=1}^l (d_i'^2 - d_i^2)}{ld^2}} = \sqrt{\frac{\dim X' - \dim X' \cap Y'}{\dim W'_j}}.
\end{equation}
Equation (3) implies that this is at most $\sqrt{4n/M}$. This completes the proof of Lemma 11.

By summing over $W_j$’s, we get
\begin{equation}
\dim X - \dim X \cap Y \leq \sqrt{\frac{4n}{M}} \sum_j \dim W_j = \sqrt{\frac{4n}{M} 2^n}
\end{equation}
and the rest is similar to the end of Section 7. This completes the proof of Theorem 3.
Appendix: Proof of the Classical Impossibility Result

PROOF OF THEOREM 2. Let $i \neq j$ and $k \neq l$. Then, there is a permutation $\pi$ such that $\pi(i) = k$, $\pi(j) = l$. By permutability, this means that there is $\sigma$ such that $\sigma(A_i) = A_j$ and $\sigma(A_j) = A_i$. Therefore, the intersections $A_i \cap A_j$ and $A_k \cap A_l$ must have the same number of elements. Thus, $|A_i \cap A_j| \leq (1 - \epsilon)|A_i| = (1 - \epsilon)2^{n-1}$ for every $i \neq j$. Equivalently, we have a bound on the set difference of $A_i$ and $A_j$: $|A_i \setminus A_j| \geq \epsilon 2^{n-1}$.

LEMMA 12. Let $M > 2/\epsilon$. There exists $x \in \{0, 1\}^n$ that belongs to $k$ sets $A_i$ for $\frac{1}{2}M \leq k \leq (1 - \frac{\epsilon}{2})M$.

PROOF. We count the number of triples $(i, j, x)$, $i, j \in \{1, \ldots, M\}, x \in \{0, 1\}^n$ such that $i \neq j$, $x \in A_i$, $x \notin A_j$.

If $|A_i \setminus A_j| \geq \epsilon 2^{n-1}$, there are at least $\epsilon 2^{n-1}$ $x$ that belong to $A_i$ but not to $A_j$. Since $A_i \setminus A_j$ is of the same size for all $i, j, i \neq j$, this must be true for every pair $i, j, i \neq j$. Since there are $M(M - 1)$ such pairs, we get at least $\epsilon M(M - 1)2^{n-1}$ triples.

On the other hand, if $x$ belongs to at least $(1 - \frac{\epsilon}{2})M$ or at most $\frac{1}{2}M$ sets $A_i$, there are at most $\frac{1}{2}M(1 - \frac{\epsilon}{2})M$ pairs $(i, j)$ such that $x \in A_i$ but $x \notin A_j$. If this is true for every of $2^n$ possible $x \in \{0, 1\}^n$, we get at most $\frac{1}{2}((1 - \frac{\epsilon}{2})M^22^n$ triples. Then, we must have

$$\epsilon M(M - 1)2^{n-1} \leq \frac{\epsilon}{2} \left(1 - \frac{\epsilon}{2}\right) M^22^n.$$

This only happens when $(M - 1) \leq (1 - \frac{\epsilon}{2})M$ which is equivalent to $\epsilon M \leq 2$. In this case, $M \leq \frac{\epsilon}{2}$. □

Let $x \in \{0, 1\}^n$ belong to $k$ sets $A_i$. Without loss of generality, assume that $x$ belongs to $A_1, A_2, \ldots, A_k$. Let $i_1, i_2, \ldots, i_k$ be $k$ different numbers from $\{1, \ldots, M\}$. Take a permutation $\pi \in S_M$ that $\pi(1) = i_1, \ldots, \pi(k) = i_k$. By the permutability requirement, there is a $\sigma \in S_N$ such that $\sigma(A_i) = A_{\pi(i)}$. This means that $\sigma(x)$ belongs to $A_{\pi(1)}, \ldots, A_{\pi(k)}$ and no other set. We have shown that for any $i_1, \ldots, i_k$, there is a $y \in \{0, 1\}^n$ that belongs to $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ and no other set. There are $\binom{M}{k}$ ways to choose $i_1, i_2, \ldots, i_k$. Therefore, there must be at least $\binom{M}{k}$ different $y \in \{0, 1\}^n$. This means that

$$\binom{M}{k} \leq 2^n.$$

Without loss of generality, we can assume that $k \leq M/2$. (Otherwise, just replace $k$ by $M - k$ and $\binom{M}{k}$ will stay the same.) We have $\binom{M}{k} \geq \left(\frac{M}{k}\right)^k$. Since $k \leq M/2$, this is at least $2^k$. Since $k \geq \frac{1}{2}M$, $2^k \geq 2^{M/2}$. Thus, we have

$$2^{\epsilon M/2} \leq 2^n.$$

This implies $\epsilon M/2 \leq n$ and $M \leq \frac{2}{\epsilon}n$. □

The large size of the subsets means that we have far more constraints than we have degrees of freedom. The combinatorial argument shows that if the requirement of full permutability is imposed, and we have a superpolynomial (in $n$) number of subsets, then the difference of every two sets must be a vanishing fraction of the
size of the sets. The two types of sets \( \{ A_b \} \) described above, however, separately achieve distinguishability and permutability.

REFERENCES


