Quantum supremacy in constant-time measurement-based computation: A unified architecture for sampling and verification

Jacob Miller, Stephen Sanders, and Akimasa Miyake

Phys. Rev. A **96**, 062320 — Published 20 December 2017

DOI: 10.1103/PhysRevA.96.062320

Quantum supremacy in constant-time measurement-based computation: A unified architecture for sampling and verification

Jacob Miller,[*] Stephen Sanders, and Akimasa Miyake[†]
Center for Quantum Information and Control, Department of Physics and Astronomy,
University of New Mexico, Albuquerque, NM 87131, USA

While quantum speed-up in solving certain decision problems by a fault-tolerant universal quantum computer has been promised, a timely research interest includes how far one can reduce the resource requirement to demonstrate a provable advantage in quantum devices without demanding quantum error correction, which is crucial for prolonging the coherence time of qubits. We propose a model device made of locally-interacting multiple qubits, designed such that simultaneous single-qubit measurements on it can output probability distributions whose average-case sampling is classically intractable, under similar assumptions as the sampling of non-interacting bosons and instantaneous quantum circuits. Notably, in contrast to these previous unitary-based realizations, our measurement-based implementation has two novel features. (i) Our implementation involves no adaptation of measurement bases, leading output probability distributions to be generated in constant time, independent of the system size. Thus, it could be implemented in principle without quantum error correction. (ii) Verifying the classical intractability of our sampling is done by changing the Pauli measurement bases only at certain output qubits. Our usage of random commuting quantum circuits in place of computationally universal circuits allows a unique unification of sampling and verification, so that they require the same physical resource requirements in contrast to the more demanding verification protocols seen elsewhere in the literature.

I. INTRODUCTION

General-purpose quantum computers hold the promise of achieving quantum speed-ups in many problems of practical importance, unmatched by any known classical methods [1–3]. While the prospect of such speed-ups is exciting, a growing realization is the extreme difficulty of achieving the levels of precision and control required for building truly scalable, fault-tolerant quantum hardware. As an intermediate step towards this goal, several recent proposals have suggested the development of special-purpose quantum devices which achieve so-called "quantum supremacy" in certain tasks [4–21]. Instead of solving general computational problems, these devices instead sample from probability distributions widely believed to be impossible to simulate efficiently using classical means. The recent explosion of proposals for such classically intractable sampling devices has begun to be matched by actual demonstrations of sampling in the laboratory [22–26], although so far still at small enough scales to allow for exact classical simulation.

An important question regarding such proposals is how far, and in what manner, we can reduce the resources required to exhibit and certify a genuine quantum advantage in sampling. The boson sampling protocol [6] shows that such quantum advantage can be achieved using simple linear optical devices and single-photon detectors. However, there are many challenges facing a realistic implementation of boson sampling, including the parallel generation of many single photons, the precise timing constraints on these photons, and the robust and accurate arrangement of the required beam splitters and phase shifters. An alternative proposal which circumvents this bottleneck is the family of instantaneous quantum polynomial-time (IQP) protocols [5, 8, 15], where sampling distributions arise from single-qubit measurements on the output of low-depth commuting quantum circuits. If a quantum device can prepare sampling distributions associated with any unitary within a circuit family, then that process would be classically intractable under reasonable conjectures from complexity theory. Furthermore, the commuting nature of these quantum circuits means that they can potentially be engineered to run in constant time, maximally avoiding the threat of environmental noise and decoherence. However, a practical issue which arises here is the extreme difficulty of engineering the arbitrary long-range interactions needed for such a constant time implementation. While these long-range interactions can be simulated by bringing distant qubits together using $SWAP$ gates before applying local entangling operations, this process would introduce a new bottleneck, the growing time required to shuttle qubits between local interaction regions. In the absence of quantum error correction, the growing influence of decoherence would quickly degrade the quality of our sampling distributions, making this straightforward implementation likely untenable for practical demonstrations of quantum supremacy.

In this paper, we show how nonadaptive measurement-based quantum computation (MQC) [27–29] can be used to sample from the distributions associated with IQP circuits, while at the same time verifying the classical intractability of this sampling process. Our protocol uses a fixed resource state preparable by a constant-depth local circuit, which is then nonadaptively measured at each

[*] jmilla@unm.edu
[†] amiyake@unm.edu

site in the Pauli X, Y, or Z bases. The setting of nonadaptive MQC allows us to replace the time complexity present in local IQP circuits (with $SWAP$ gates) by a spatial overhead in our resource state, which results in a protocol with constant runtime and local interactions. The cost of this nonadaptivity is a fundamental randomness in the distributions prepared by our protocol, arising from random MQC byproduct operators. This leads each sample in our protocol to be obtained with high probability from a different sampling distribution every time. Surprisingly, we show that this inherent randomness has no impact on the hardness of our protocol, which remains classically intractable under the same assumptions as in [8]. What's more, we show that these random byproduct operators actually simplify our implementation relative to a direct circuit-based counterpart, revealing an inherent advantage of MQC for quantum sampling protocols. We further show that by simply changing the single-qubit Pauli measurements used in the final step of our protocol to obtain sampling statistics, we can instead rigorously verify the classical intractability of our sampling. Our verification scheme is inspired by the ground state certification protocol of [30], but uses the special form of our IQP sampling distributions to replace the nonlocal operations required for general Hamiltonian measurements with measurements of single-qubit Pauli operators. This lets us switch between sampling and measurement by a simple change in single-qubit measurement bases, allowing our procedure to achieve a robust demonstration of quantum supremacy capable of efficiently detecting any errors which could potentially harm the correctness of our sampling distributions.

Our protocol is closely related to that of [8], as it constitutes a faithful translation of their circuit-based IQP sampling into the context of MQC. However, we show that this translation itself contains several surprises, ultimately revolving around the nontrivial interface of MQC byproduct operators with classically intractable sampling. At first glance, our protocol has much in common with [13, 19], which also use nonadaptive MQC to perform classically intractable sampling and verification. Upon further investigation however, the different protocols are seen to utilize completely different mechanisms for demonstrating quantum supremacy, which allow for substantial differences in behavior. While using a more involved resource state than the Ising-like states of [13, 19], the design of our protocol allows a unique duality between sampling and verification, in that both require the same physical resources and are switchable by a mere change of single-qubit Pauli measurement bases on an n-qubit output state. This feature fundamentally depends upon the convenient mathematical nature of our IQP sampling distributions, and cannot be straightforwardly reproduced within the setting of sampling from random universal circuits such as [13, 19].

In Section II, we review the relevant theory behind IQP sampling, verification, and MQC. In Section III we present our protocol for preparing, sampling from, and verifying different classically intractable sampling distributions using Pauli measurements on a model resource state $|\Psi_{\text{Prep}}\rangle$. In Section IV we comment on the features unique to our protocol, and outline future directions for our work. A brief comparison of our proposal to other proposals within the rapidly growing field of classically intractable sampling can be found in Appendix A, with detailed proofs of the classical intractability and verification of our sampling protocol found in Appendices B, C and D.

II. BACKGROUND

A. IQP and Boolean Functions

In the IQP sampling protocols of [5, 8, 15], a sampling state $|\psi_f\rangle = U_f|+\rangle^{\otimes n}$ is first prepared using an n-qubit diagonal unitary circuit U_f, and is then measured everywhere in the Pauli X basis to obtain a random outcome $|\mathbf{s}_X\rangle = H^{\otimes n}|\mathbf{s}\rangle$. In the above, $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ denotes the $+1$ eigenstate of X, H the single-qubit Hadamard operator, $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ a bit string of length n, and $|\mathbf{s}\rangle$ the corresponding Z basis product state. If U_f is chosen from an appropriate family of diagonal unitaries, then [5] shows that the act of sampling from the distribution $D_f(\mathbf{s}) = |\langle \mathbf{s}_X|\psi_f\rangle|^2$ is impossible to perform in polynomial time using a classical computer, assuming the widely conjectured non-collapse of the polynomial hierarchy of complexity theory [31, 32]. More generally, we use the phrase classically intractable sampling to mean any sampling protocol which shares this property of being impossible to simulate classically (given the non-collapse of the polynomial hierarchy), possibly in the presence of some allowable error and under the assumed truth of additional mathematical conjectures.

We now choose the n-qubit unitary gates U_f above to be parameterized by n-bit binary functions $f : GF(2)^n \to GF(2)$, where $GF(2) \simeq \{0,1\}$ denotes the finite field of binary numbers. The functions f set the eigenvalues of U_f as

$$U_f = \sum_{\mathbf{x}\in GF(2)^n} (-1)^{f(\mathbf{x})}|\mathbf{x}\rangle\langle\mathbf{x}|, \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. When applied to $|+\rangle^{\otimes n}$, this results in the sampling state

$$|\psi_f\rangle = 2^{-n/2} \sum_{\mathbf{x}\in GF(2)^n} (-1)^{f(\mathbf{x})}|\mathbf{x}\rangle. \tag{2}$$

We can alternately describe $|\psi_f\rangle$ as the unique state satisfying the n (nonlocal) stabilizer relations $h_f^{(i)}\,|\psi_f\rangle = (+1)\,|\psi_f\rangle$ for $1 \le i \le n$, where

$$\begin{aligned} h_f^{(i)} &= U_f X_i U_f^\dagger \\ &= X_i \sum_{\mathbf{x}\in GF(2)^n} (-1)^{\partial_i f(\mathbf{x})}|\mathbf{x}\rangle\langle\mathbf{x}|, \end{aligned} \tag{3}$$

and the polynomial $\partial_i f$ is equal to the difference

$$\partial_i f(\mathbf{x}) = f(x_1, \ldots, x_i+1, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n). \tag{4}$$

Because addition in $GF(2)$ is modulo 2, it is easy to verify that $\partial_i f(\mathbf{x})$ is always independent of the value of x_i.

We now restrict our binary functions to be cubic polynomials, so that $f(\mathbf{x})$ can be written in the form

$$f(\mathbf{x}) = \sum_{1 \leq i<j<k \leq n} a_{ijk} x_i x_j x_k + \sum_{1 \leq i<j \leq n} b_{ij} x_i x_j + \sum_{1 \leq i \leq n} c_i x_i, \tag{5}$$

for some binary coefficients a_{ijk}, b_{ij}, and c_i. These are generated by linear, quadratic, and cubic monomials, whose associated diagonal unitary gates are $U_{x_i} = Z_i$, $U_{x_i x_j} = CZ_{ij}$ (controlled-Z), and $U_{x_i x_j x_k} = CCZ_{ijk}$ (controlled-controlled-Z). More explicitly, the gates Z_i, CZ_{ij}, and CCZ_{ijk} are defined by their action on qubits i, j, and k respectively as $Z_i|x_i\rangle = (-1)^{x_i}|x_i\rangle$, $CZ_{ij}|x_i, x_j\rangle = (-1)^{x_i x_j}|x_i, x_j\rangle$, and $CCZ_{ijk}|x_i, x_j, x_k\rangle = (-1)^{x_i x_j x_k}|x_i, x_j, x_k\rangle$. In the following, any references to polynomials will be understood to refer specifically to binary polynomials. We will use $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ to denote homogeneous polynomials, for which the only nonzero coefficients are of the form a_{ijk}, b_{ij}, or c_i, respectively. Similarly, $\mathbf{b}+\mathbf{c}$ and $\mathbf{a}+\mathbf{b}$ will denote polynomials for which all $a_{ijk} = 0$ or all $c_i = 0$, respectively.

It will be convenient in the following to interpret n-bit vectors $\mathbf{s}$ as linear polynomials of n variables, which act as

$$\mathbf{s}(\mathbf{x}) = \sum_{i=1}^{n} s_i x_i. \tag{6}$$

This is useful in giving the probability of different sampling outcomes, as the probability of obtaining any given $|\mathbf{s}_X\rangle$ when $|\psi_f\rangle$ is measured in the X product basis is

$$\begin{aligned} D_f(\mathbf{s}) &= \left|\langle \mathbf{s}_X|\psi_f\rangle\right|^2 \\ &= \left|2^{-n} \sum_{\mathbf{x} \in GF(2)^n} (-1)^{f(\mathbf{x})+\mathbf{s}(\mathbf{x})}\right|^2 \\ &= \mathrm{ngap}^2(f+\mathbf{s}). \end{aligned} \tag{7}$$

$\mathrm{ngap}^2(f)$ refers here to the square of $\mathrm{ngap}(f)$, the (signed) difference between the fraction of inputs yielding $f(\mathbf{x}) = 0$ and $f(\mathbf{x}) = 1$. $\mathrm{ngap}(f)$ is known to be #P-hard to compute for arbitrary cubic polynomials f [33], and we will see that this hardness underlies the classical intractability of our sampling protocol.

B. Classically Intractable Sampling and Verification

It is shown in [8] that estimating the quantity $\mathrm{ngap}^2(f)$ up to $\frac{1}{4}$ multiplicative error, so that $|\mathrm{ngap}^2_{\mathrm{Est}}(f) - \mathrm{ngap}^2(f)| \leq \frac{1}{4}\mathrm{ngap}^2(f)$ for arbitrary cubic polynomials f, is #P-hard, mirroring the difficulty of computing $\mathrm{ngap}(f)$. This hardness leads to a similar finding as in [5], that exactly sampling from the cubic polynomial distributions D_f defined in Eq. (7) is classically intractable. In particular, assuming the existence of a classical randomized algorithm which can efficiently sample from any of the distributions D_f lets a technique called Stockmeyer approximate counting [34] be used to estimate the probabilities $D_f(\mathbf{s})$ up to $\frac{1}{4}$ multiplicative error, and thus to solve arbitrary #P problems. While Stockmeyer counting is an unphysical process which cannot be implemented with realistic classical or quantum computers, it can be carried out at a finite level of the polynomial hierarchy, and the hardness of #P problems for this hierarchy then leads to its collapse. Details of this process can be found in Appendix C. On the other hand, we have seen that these distributions appear naturally as the output distributions of the IQP sampling protocol described above, which allows us to interpret a concrete implementation of this protocol as a provable demonstration of "quantum supremacy".

While straightforward and conceptually compelling, a major limitation of the above result is the impossibility of verifying that any realistic quantum protocol is sampling from *exactly* the ideal distribution D_f [35]. In order to demonstrate quantum supremacy in a more realistic setting, an alternate proof is given in [8] which shows the classical intractability of sampling from any distribution Q_f which is variationally close to D_f. Variationally close means here that the statistical distance between Q_f and D_f is bounded by a constant η_0, so that

$$\left|Q_f - D_f\right|_1 = \sum_{\mathbf{s} \in GF(2)^n} |Q_f(\mathbf{s}) - D_f(\mathbf{s})| \leq \eta_0, \tag{8}$$

In [8] a value of $\eta_0 \leq \frac{1}{192}$ was shown to be sufficient for classically intractable sampling, which in Appendix C we show can be relaxed to $\eta_0 \leq \frac{1}{86}$ (although both values rely on the particular value of ϵ_0 appearing in Conjecture 1 below). This result is appealing from a practical standpoint, as the quantity $\left|Q_f - D_f\right|_1$ can be efficiently estimated in experiments involving quantum sampling distributions.

On the other hand, the above "average-case" sampling result relies upon one additional complexity theoretic conjecture:

Conjecture 1 (Average-Case Hardness of $\mathrm{ngap}^2(f)$). *Let f be an arbitrary cubic polynomial of the form given in Eq. (5). Then it is #P-hard to efficiently calculate an estimate* $\mathrm{ngap}^2_{\mathrm{Est}}(f)$ *of* $\mathrm{ngap}^2(f)$ *for which* $|\mathrm{ngap}^2_{\mathrm{Est}}(f) - \mathrm{ngap}^2(f)| \leq \frac{1}{4}\mathrm{ngap}^2(f)$, *on at least* $1 - \epsilon_0 = \frac{1}{24}$ *of polynomials f.*

Intuitively, this conjecture states that even when our estimates $\mathrm{ngap}^2_{\mathrm{Est}}(f)$ are allowed to fail with some finite probability ϵ_0, corresponding to realistic errors in our sampling distributions Q_f, the problem of estimating $\mathrm{ngap}^2(f)$ on the remaining instances is still #P-hard.

While this reliance on an additional unproven conjecture isn't desirable, an analogous conjecture is required for every known average-case classically intractable sampling result, and thus isn't any special demerit of [8].

The techniques of [30] can be used to efficiently verify the condition $\left|Q_f - D_f\right|_1 \leq \eta_0$ when Q_f arises from measurements on experimentally prepared quantum sampling states ρ_f, which approximate our intended $|\psi_f\rangle$. Given ρ_f, we can perform measurements of the nonlocal Hermitian stabilizers $h_f^{(i)}$ defined in Eq. (3), which will always yield the outcome +1 in the ideal case where $\rho_f = |\psi_f\rangle\langle\psi_f|$. In more general cases, a sufficiently accurate empirical estimate of these n observables $h_f^{(i)}$ can be converted into a bound on the statistical distance between the distributions Q_f and D_f. If the average $\langle h_f^{(i)}\rangle$ is sufficiently close to +1 so as to guarantee $\left|Q_f - D_f\right|_1 \leq \eta_0$, then we can confidently conclude that our quantum protocol is performing classically intractable sampling. We will soon show that the nonlocal measurements of $h_f^{(i)}$ can actually be entirely replaced with single-qubit X and Z measurements, which allows this verification to be done within the setting of MQC.

C. Measurement-Based Quantum Computation

MQC is a means of carrying out computation using only single-qubit measurements on a fixed many-body resource state. In this framework, the choice of measurements made on local regions of our resource state determines logical operations which are applied to encoded logical qubits, while simultaneously teleporting these qubits to adjacent unmeasured sites. The randomness of quantum measurement leads the outcomes of these measurements to determine a so-called byproduct operator, which acts as a random correction to the overall logical operation. For example, in Figure 1a we show the standard protocol for teleporting one logical qubit within the MQC quantum wire known as the 1D cluster state. Given two successive X measurements with outcomes $|t_{1,X}\rangle$ and $|t_{2,X}\rangle$, the resultant logical operation is $U_X(t_1, t_2) = X^{t_2} Z^{t_1}$, showing the intended logical unitary to be the identity and the byproduct operator to be a random Pauli $X^{t_2} Z^{t_1}$. In Figure 1b we show a gadget for performing the two-qubit $SWAP$ operation on logical qubits, for which the byproduct operator is a random two-qubit Pauli operator. In both of these examples, the collection of operators appearing as byproducts for arbitrary measurement outcomes form a closed group (up to global phase) of finite size, referred to as a byproduct group.

An MQC protocol is said to be adaptive if the choice of measurement in some region of our resource state depends on the outcome of measurements made in another region. Adaptation can be seen as a means of ensuring that the byproduct group associated with a large computation remains finite (for example, contained within the n-qubit Pauli group), whereas the use of nonadaptive MQC with arbitrary single-qubit measurements will generally lead to a byproduct group of unbounded size. On the other hand, nonadaptive MQC computations can always be implemented in constant time by performing all measurements simultaneously, a serious advantage in the absence of quantum error correction. Within the usual scheme for universal MQC using resource states built from CZ gates, nonadaptive single-qubit Pauli measurements are associated with byproduct groups formed from Pauli operators, and implement logical operations contained within the Clifford group. The Clifford group is defined as those unitaries U which preserve the Pauli group under conjugation, so that $UPU^\dagger$ is a product of Pauli operators whenever P is. The evolution of Pauli eigenstates under the Clifford group is known to be efficiently simulable using classical means [36], which means that non-Clifford operations are necessary for demonstrating quantum supremacy.

In Figure 1c, we give an example of an MQC gadget which implements a non-Clifford CCZ gate when nonadaptive Pauli measurements are applied. This gadget, which will be utilized in our classically intractable sampling protocol below, is itself formed from non-Clifford CCZ gates, and has a byproduct group containing non-Pauli CZ gates. A similar gadget was shown in [37] to enable universal MQC using only Pauli measurements, but with adaptation of measurement bases so as to avoid a byproduct group of unbounded size. In our MQC sampling protocol below, we will show that restricting our logical operations to those generating sub-universal quantum computation will allow us to avoid this use of adaptation, while still maintaining a byproduct group of finite size. In fact, we will find that this non-Pauli byproduct group actually leads to a simplification in our protocol relative to circuit-based counterparts.

III. MQC PROTOCOL FOR CLASSICALLY INTRACTABLE SAMPLING

Our MQC implementation of the classically intractable sampling protocol of [8] uses nonadaptive Pauli measurements to prepare, sample from, and verify the n-qubit sampling states $|\psi_f\rangle$ described above, for arbitrary cubic polynomials f. Our protocol uses a 2D resource state $|\Psi_{\text{Prep}}\rangle$ which is capable of preparing any sampling state $|\psi_f\rangle$ using only single-qubit Pauli measurements. $|\Psi_{\text{Prep}}\rangle$ is constructed from the teleportation, $SWAP$, and CCZ gadgets described in Section II C, which are configured to implement any of the IQP circuits $U_{\mathbf{a}}$ associated with arbitrary homogeneous cubic polynomials $\mathbf{a}$. The choice of $\mathbf{a}$ is determined by the choice of Pauli measurement basis applied to each CCZ gadget in $|\Psi_{\text{Prep}}\rangle$. By virtue of the byproducts arising from our nonadaptive MQC implementation, our output sampling states end up being random $|\psi_f\rangle$ where $f = \mathbf{a}+\mathbf{b}+\mathbf{c}$ is a sum of the intended $\mathbf{a}$, along with random quadratic and linear polynomials

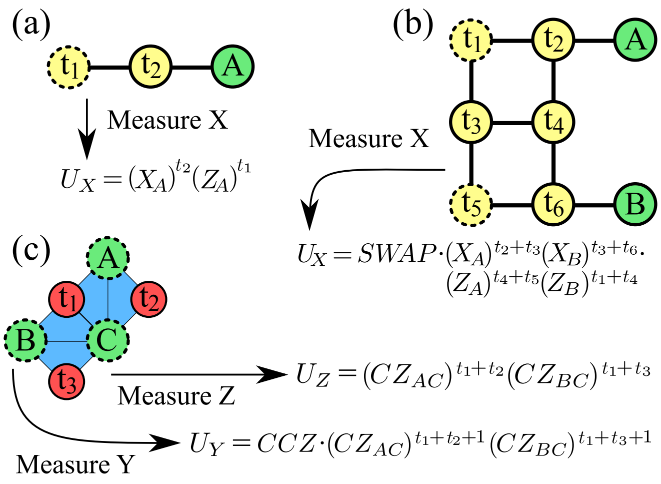
(a)
t1
t2
A
Measure X
$U_X = (X_A)^{t_2}(Z_A)^{t_1}$
(b)
t1
t2
A
t3
t4
t5
t6
B
Measure X
$U_X = SWAP{\cdot}(X_A)^{t_2+t_3}(X_B)^{t_3+t_6}{\cdot}(Z_A)^{t_4+t_5}(Z_B)^{t_1+t_4}$
(c)
A
t1
t2
B
C
t3
Measure Z
$U_Z = (CZ_{AC})^{t_1+t_2}(CZ_{BC})^{t_1+t_3}$
Measure Y
$U_Y = CCZ{\cdot}(CZ_{AC})^{t_1+t_2+1}(CZ_{BC})^{t_1+t_3+1}$

FIG. 1. The MQC gadgets utilized in our protocol. We describe the formation circuit and outcome-dependent logical operations for each when sites are measured everywhere in some single-qubit Pauli basis. Initial states on input sites (dotted) are teleported to output sites (green online or gray in print grayscale) and acted on by characteristic logical operations and measurement-dependent random byproduct operators. These outputs are identified with the inputs of future gadgets. The relationship between the formation circuits shown and the logical operations implemented is given by contracting with the appropriate measurement outcomes, which additionally contributes a scalar factor of $\frac{1}{\sqrt{2}}$ per measurement (not shown). Mathematically, this leads our measurement outcomes s_i to occur uniformly randomly. (a) 1D cluster state wire of length 2, where solid lines indicate CZ formation unitaries. Measuring X on two sites implements the identity, with a uniformly random Pauli byproduct group. (b) Planar MQC gadget for implementing nonplanar wire crossings. Measuring X on 6 sites implements $SWAP$, with a byproduct group of uniformly random two-qubit Pauli operators. (c) Non-Clifford gadget for conditional CCZ, where triangles indicate CCZ gates used to form the gadget. Measuring Y on 3 non-logical control sites (dark red, smaller circles) gives CCZ on sites A, B, and C, whereas measuring Z on these sites instead gives the identity. In both cases, the teleportation is trivial (output and input sites coincide), while the byproduct group is a product of uniformly random CZ's between A and C, and between B and C.

b and **c**. Owing to this randomness in $\mathbf{b}+\mathbf{c}$, we are unable to deterministically prepare any fixed sampling state $|\psi_f\rangle$. Despite this fundamental indeterminism, we will show how the act of sampling from randomly prepared $|\psi_f\rangle$ with X measurements at the final stage of our protocol remains classically intractable, even in the presence of realistic noise which leads our output sampling distributions to be some imperfect Q_f. We state the classical intractability of our protocol, and the precise conditions which guarantee this, as Theorem 1.

Theorem 1. *Assume the validity of Conjecture 1 and the non-collapse of the polynomial hierarchy. If the distributions $Q_f(\mathbf{s})$ arising from our MQC sampling protocol are close on average to the distributions $D_f(\mathbf{s})$ defined in Eq. (7), meaning that the average ℓ_1 norm over all f meets the experimental threshold $\langle |Q_f - D_f|_1 \rangle_f \leq \eta_0 = \frac{1}{86}$, then our protocol is impossible to efficiently simulate using a classical computer, i.e. is classically intractable.*

Our protocol for classically intractable sampling is divided into two stages: preparation of the random sampling state $|\psi_f\rangle$ and sampling/verification measurements on $|\psi_f\rangle$ (see Figure 2). In the preparation stage, we use $m = O(n^4)$ single-qubit measurements of Pauli X, Y, and Z on $|\Psi_{\rm Prep}\rangle$ with outcomes $\mathbf{t} = (t_1, t_2, \ldots, t_m)$ to prepare the n-qubit state $|\psi_{f(\mathbf{t})}\rangle$ associated with a $\mathbf{t}$-dependent polynomial $f(\mathbf{t}) = \mathbf{a} + \mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t})$. These measurements are chosen to implement the unitary $U_{\mathbf{a}}$ by means of a depth $O(n^3)$ quantum circuit built from local CCZ and $SWAP$ gates. The CCZ gates in this ideal circuit are applied conditionally as $(CCZ)^{a_{ijk}}$, depending on the coefficients of $\mathbf{a}$, with teleportation and $SWAP$ gates used before each application to move qubits i, j, and k into the same region. The application of these conditional CCZ's is structured within three nested levels of iteration, which together apply all $\binom{n}{3}$ three-body terms in the lexicographic order of the triples (i, j, k), where $i < j < k$. Loop I, the lowest level of iteration, involves fixing qubits i and j in a designated interaction region, then successively cycling the remaining qubits $k > j$ through this region. $(CCZ)^{a_{ijk}}$ is applied in turn to each triple, until all triples (i, j, k) with fixed i and j have been processed in this manner. Loop II, the next level of iteration, involves successively replacing qubit j by qubit $j+1$, then repeating Loop I for all qubits $k > j+1$ until all triples (i, j, k) with fixed i have been processed. Finally, Loop III involves successively replacing qubit i by qubit $i+1$, in the process shifting the location of the interaction region, and repeating Loop II for all qubits $j, k > i+1$ until $(CCZ)^{a_{ijk}}$ has been applied to all triples of qubits. The resulting unitary operation is clearly $U_{\mathbf{a}}$.

While the simple circuit described above is only capable of producing sampling states $|\psi_{\mathbf{a}}\rangle$ associated with homogeneous cubic $\mathbf{a}$, our MQC implementation utilizes random byproduct operators to implement the remaining quadratic and linear terms required for the preparation of arbitrary $|\psi_f\rangle$. This reveals a simplification within nonadaptive MQC compared to a direct circuit-based counterpart, which would require additional CZ and Z gates to implement U_f for arbitrary f. Each of the conditional operations $(CCZ)^{a_{ijk}}$ is implemented using the CCZ gadget shown in Figure 1c, which is measured in Y if $a_{ijk} = 1$ and Z otherwise. For either choice of measurement, the non-Clifford nature of these gadgets leads the resultant byproduct operators to consist of non-Pauli CZ gates, which generate random quadratic terms in the output $|\psi_f\rangle$. Because our logic gates and byproduct operators are made up of X and the diagonal Z, CZ, and CCZ gates, which together form a closed (non-universal) gate set under multiplication, the byproduct group associated with our computation will always remain finite. This is in contrast to the byproduct group appearing in MQC

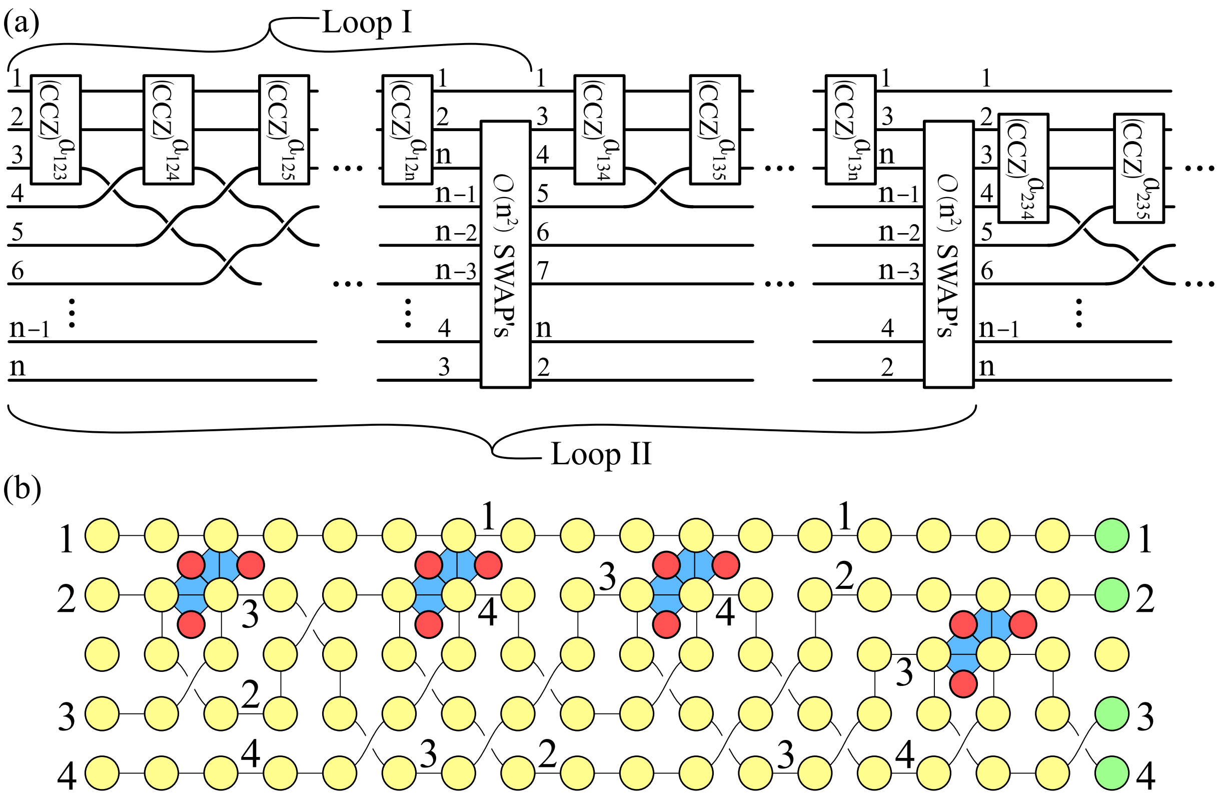

FIG. 2. An overview of our constant-time MQC protocol for implementing the unitary $U_f = U_{\mathbf{b}+\mathbf{c}}U_{\mathbf{a}}$ which prepares the sampling state $|\psi_f\rangle$. Our intended logical operation is $U_{\mathbf{a}}$, while $U_{\mathbf{b}+\mathbf{c}}$ is a byproduct contribution containing uniformly random $\mathbf{b}$ and $\mathbf{c}$. (a) Circuit diagram for $U_{\mathbf{a}}$, which is formed from several repeating loops. In Loop I, qubits i_0 and j_0 remain fixed and all qubits $k > i_0, j_0$ are sequentially cycled past i_0 and j_0 and acted on by a conditional three-body gate $(CCZ_{i_0j_0k})^{a_{i_0j_0k}}$ depending on the binary coefficient $a_{i_0j_0k}$ in f. The order of these qubits is reversed after Loop I, which is undone by a sequence of $SWAP$'s with circuit depth $O(n)$. Loop II then involves replacing qubit j_0 by j_0+1, and repeating Loop I for all triples (i_0, j_0+1, k), where $k > i_0, j_0+1$. Loop II continues cycling qubit j and applying Loop I until all triples (i_0, j, k) have been addressed. Loop III (not shown) then involves replacing qubit i_0 by i_0+1, and repeating Loop II for all triples (i_0+1, j, k). At the completion of Loop III, we have addressed all triples of qubits within circuit depth $O(n^3)$, producing the output state $|\psi_{\mathbf{a}}\rangle$. (b) A concrete example of how the above protocol is implemented in MQC using our resource state $|\Psi_{\text{Prep}}\rangle$, for $n = 4$. 1D cluster state wires let us teleport information between non-Clifford gadgets, which apply the logical gate $(CCZ)^{a_{ijk}}$ via an a_{ijk}-dependent choice of Y or Z measurement on control sites (dark red, smaller circles). While our state is drawn with nonplanar wire crossings, these are simulated using the planar $SWAP$ gadgets in Figure 1b. Measuring all preparation sites simultaneously prepares a random n-qubit state $|\psi_f\rangle$ on the output sites (on right, green online or gray in print grayscale), where $f = \mathbf{a}+\mathbf{b}+\mathbf{c}$ contains a deterministic $\mathbf{a}$ set by the measurement bases and a uniformly random $\mathbf{b}+\mathbf{c}$ arising from random byproduct operators. The final n-qubit measurement is chosen to randomly implement sampling via all X measurements, or verification via a mixture of X and Z measurements.

implementations of random circuit quantum supremacy protocols, such as [13, 19], which grows unboundedly.

The CCZ gadgets used in our protocol are embedded in regular intervals in $|\Psi_{\text{Prep}}\rangle$, and are then connected together using 1D cluster wires and $SWAP$ gadgets, which simulate the movement of qubits utilized in our ideal quantum circuit described above. These cluster wires and $SWAP$ gadgets are always measured in X, which leads to a product of random Pauli X and Z byproduct operators. The Z byproducts eventually end up generating random linear terms in the output state $|\psi_f\rangle$, while the X byproducts can be commuted backwards in our circuit, to eventually be annihilated on the initial $|+\rangle^{\otimes n}$ which our logical quantum circuit is applied to. This commutation of X byproduct operators induces conditional $(CZ)^{a_{ijk}}$ and $(Z)^{a_{ijk}}$ byproduct operators arising from prior CCZ gadgets, which results in additional randomness in the overall byproduct group. Despite this seeming complexity in the distribution of byproduct operators, we prove in Appendix B that the random outcomes $\mathbf{t}$ of preparation measurements on $|\Psi_{\text{Prep}}\rangle$ lead the random quadratic and linear terms in the polynomial $f = \mathbf{a}+\mathbf{b}+\mathbf{c}$ associated with $|\psi_f\rangle$ to be uniformly random, simplifying our analysis.

In the second stage of our protocol we apply a final series of n single-qubit Pauli measurements to our output state which, while ideally equal to $|\psi_f\rangle\langle\psi_f|$, will realistically be some mixed state ρ_f. The choice of single-qubit measurement bases depends on whether we are implementing sampling or verification, which can be chosen randomly with $\frac{1}{2}$ probability. During sampling, we simply measure all qubits in the X basis to generate a sample from the distribution $Q_f(\mathbf{s}) = \mathrm{Tr}(\rho_f |\mathbf{s}_X\rangle\langle \mathbf{s}_X|)$, exactly as described in Section II B. Although the randomness in the f associated with ρ_f means that we will almost certainly obtain each sample from a different distribution Q_f, our MQC sampling protocol remains classically intractable nonetheless. To prove this classical intractability, we can treat the overall process of preparing a random ρ_f and then sampling an outcome s as itself a sampling process with probability $\mathrm{Pr}_{\mathbf{a}}(\mathbf{b}+\mathbf{c}, \mathbf{s})$. Given this description, and our knowledge of the complete randomness of the byproduct contributions $\mathbf{b}+\mathbf{c}$, Stockmeyer approximate counting can then be used to estimate $Q_f(\mathbf{s})$ as a conditional probability which is directly proportional to $\mathrm{Pr}_{\mathbf{a}}(\mathbf{b}+\mathbf{c}, \mathbf{s})$. This suffices to proves Theorem 1 using the same arguments as in other classically intractable sampling proposals, the details of which are given in Appendix C.

If we choose to perform verification instead of sampling, then we measure all qubits in the Z basis, except for a random qubit i which is measured in X. The outcome of this measurement $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is then fed into a parity function $\pi_f^{(i)}(\mathbf{v}) = \partial_i f(\mathbf{v}) + v_i$, where $\partial_i f(\mathbf{v})$ is defined in Eq. (4). This process results in an output value of 0 or 1, which we show in Appendix D gives the same information as a measurement of the nonlocal stabilizer $h_f^{(i)}$ described in Eq. (3), with outcome $(-1)^{\pi_f^{(i)}(\mathbf{v})}$. Because of our ability to characterize the closeness of ρ_f to $|\psi_f\rangle\langle\psi_f|$ using measurements of $h_f^{(i)}$, this means that we can interpret $\pi_f^{(i)}(\mathbf{v}) = 0$ as a successful verification measurement, and $\pi_f^{(i)}(\mathbf{v}) = 1$ as a deviation of ρ_f from our intended $|\psi_f\rangle$. By obtaining many samples of $\pi_f^{(i)}(\mathbf{v})$ for random i, $\mathbf{v}$, and ρ_f, the resultant estimate of $\langle \pi_f^{(i)} \rangle$ lets us guarantee the classical intractability of our MQC sampling protocol to any desired statistical significance using only $O(n^2)$ rounds of verification measurements, as stated in Theorem 2.

Theorem 2. *Suppose that the empirical average of our parity function after μn^2 verification measurements satisfies $\langle \pi_f^{(i)}(\mathbf{v}) \rangle_{\mathbf{v},i,f} \leq \frac{\eta_0^2}{n}$, for the η_0 appearing in Theorem 1. Then we can conclude with probability $p \geq 1 - e^{-O(\mu^2)}$ that our sampling distributions Q_f satisfy the assumptions of Theorem 1, and thus generate classically intractable sampling.*

We give a detailed proof of Theorem 2 in Appendix D. We should mention that another potential means of verifying the classical intractability of our sampling protocol would have been to directly measure the $O(n^4)$ local stabilizers of our resource state $|\Psi_{\mathrm{Prep}}\rangle$, analogous to the technique used in [13, 19]. The idea behind this verification scheme is that, if we guarantee our MQC resource state to be the ideal $|\Psi_{\mathrm{Prep}}\rangle$, then performing our prescribed Pauli measurements should always generate the ideal sampling states $|\psi_f\rangle$. Unfortunately, this resource state verification scheme doesn't detect errors occurring during preparation measurements, so that even when given an ideal MQC resource state, measurement imperfections during state preparation will still lead to logical errors which harm our output sampling state ρ_f. In order for this verification scheme to rigorously guarantee the classical intractability of sampling in our setting, the single-qubit error rates for measurement must be less than $O(n^{-4})$, whereas our verification technique only needs errors rates of $O(n^{-1})$. Since this latter rate is the maximum allowed for any kind of sampling to maintain a constant variational error, this shows our verification scheme to be optimal with regards to its soundness under measurement imperfections. The techniques used to achieve these more favorable allowed error rates fundamentally rely on our use of Conjecture 1, and cannot be directly transferred to other sampling settings such as [13, 19].

IV. OUTLOOK

We have demonstrated the use of MQC to perform classically intractable sampling and verification in a unified manner, with identical resource requirements for each task. This shows that verifying the hardness of a quantum sampling protocol doesn't need to be any harder than the actual sampling, and in certain architectures comes essentially for free. This contrasts sharply with many existing quantum supremacy proposals[6, 12, 16, 20], for which verifying the non-classical nature of sampling is significantly harder than the sampling itself, likely requiring exponential computational resources to ensure correctness. By using nonadaptive MQC to drive our protocol, we have furthermore allowed both sampling and verification to be carried out in constant time, which minimizes the effect of environmental decoherence, and potentially allows us to avoid the use of quantum error correction.

As an outlook, we expect that a hybrid MQC sampling platform combining the simple physical implementation of [13] or [19] with the convenient theoretical analysis and flexibility available here would represent an extremely appealing framework for implementing classically intractable sampling. In particular, a sampling protocol using nonadaptive MQC with non-Clifford $\sqrt{CZ}$ gadgets embedded in a 2D brickwork-type lattice could potentially demonstrate quantum supremacy in constant time using only $O(n\log(n))$ qubits, and with entirely local interactions. Such a protocol would implement the "sparse" IQP circuits appearing in [15], which require only $O(n\log(n))$ two-body interactions. While this can be implemented in our framework using a 2D lattice of $O(n^2\log(n))$ qubits which generalizes our $|\Psi_{\mathrm{Prep}}\rangle$, the possibility of reducing resource requirements further, potentially to $O(n\log(n))$ qubits, would require using local complementation operations on graph states. As these operations can quickly generate long-range entanglement using only local Y basis measurements, we consider such capabilities to represent a unique feature of MQC which are well-suited to reproducing the long-range, low-depth quantum circuits often utilized for quantum sampling.

V. ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grants PHY-1314955, PHY-1521016, and PHY-1620651.

[1] P.W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Sci. Statist. Comput. **26**, 1484 (1997).

[2] L.K. Grover, *Quantum Mechanics Helps in Searching for a Needle in a Haystack*, Phys. Rev. Lett. **79**, 325 (1997).

[3] D. Deutsch and R. Jozsa, *Rapid Solution of Problems by Quantum Computation*, Proc. R. Soc. London A **439**, 553 (1992).

[4] B.M. Terhal, D.P. DiVincenzo, *Adaptive Quantum Computation, Constant Depth Quantum Circuits and Arthur-Merlin Games*, Quant. Inf. Comp. **4**, 134–145 (2004).

[5] M.J. Bremner, R. Jozsa, D.J. Shepherd, *Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy*, Proc. R. Soc. A **467**, 459–472 (2011).

[6] S. Aaronson and A. Arkhipov, *The computational complexity of linear optics*, Theory of Computing, 9(4):143252, (2013).

[7] T. Morimae, K. Fujii, and J.F. Fitzsimons, *On the hardness of classically simulating the one clean qubit model*, Phys. Rev. Lett. **112**, 130502 (2014).

[8] M.J. Bremner, A. Montanaro, and D.J. Shepherd, *Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations*, Phys. Rev. Lett. **117**, 080501 (2016).

[9] S. Rahimi-Keshari, T.C. Ralph, C.M. Caves, *Sufficient Conditions for Efficient Classical Simulation of Quantum Optics*, Phys. Rev. X **6**, 021039 (2016).

[10] E. Farhi and A.W. Harrow, *Quantum Supremacy through the Quantum Approximate Optimization Algorithm*, arXiv:1602.07674, (2016).

[11] P.P. Rohde, D.W. Berry, K.R. Motes, J.P. Dowling, *A Quantum Optics Argument for the #P-hardness of a Class of Multidimensional Integrals*, arXiv:1607.04960, (2016).

[12] S. Boixo *et al.*, *Characterizing Quantum Supremacy in Near-Term Devices*, arXiv:1608.00263, (2016).

[13] 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).

[14] K. Fujii, *Noise Threshold of Quantum Supremacy*, arXiv:1610.03632, (2016).

[15] M. Bremner, A. Montanaro, and D. Shepherd, *Achieving quantum supremacy with sparse and noisy commuting quantum computations*, arXiv:1610.01808, (2016).

[16] B. Fefferman, M. Foss-Feig, and A.V. Gorshkov, *Exact sampling hardness of Ising spin models*, arXiv:1701.03167, (2017).

[17] F. Shahandeh, A.P. Lund, T.C. Ralph, *Quantum Correlations in Nonlocal BosonSampling*, arXiv:1702.02156, (2017).

[18] A.P. Lund, M.J. Bremner, T.C. Ralph, *Quantum Sampling Problems, BosonSampling and Quantum Supremacy*, arXiv:1702.03061, (2017).

[19] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, *Architectures for Quantum Simulation Showing Quantum Supremacy*, arXiv:1703.00466, (2017).

[20] A. Deshpande, B. Fefferman, M. Foss-Feig, and A.V. Gorshkov, *Complexity of sampling as an order parameter*, arXiv:1703.05332, (2017).

[21] T. Kapourniotis and A. Datta, *Nonadaptive fault-tolerant verification of quantum supremacy with noise*, arXiv:1703.09568, (2017).

[22] J.B. Spring *et al.*, *Boson sampling on a photonic chip*, Science **339**, 798–801 (2013).

[23] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Experimental boson sampling*, Nature Photon. **7**, 540–544 (2013).

[24] Crespi, A. *et al.*, *Integrated multimode interferometers with arbitrary designs for photonic boson sampling*, Nature Photon. **7**, 545–549 (2013).

[25] M.A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. Ralph, and A.G. White, *Photonic boson sampling in a tunable circuit*, Science **339**, 794–798 (2013).

[26] H. Wang *et al.*, *Multi-photon boson-sampling machines beating early classical computers*, arXiv:1612.06956, (2016).

[27] R. Raussendorf and H.J. Briegel, *A One-Way Quantum Computer*, Phys. Rev. Lett. **86**, 5188 (2001).

[28] R. Jozsa, *An Introduction to Measurement Based Quantum Computation*, arXiv:quant-ph/0508124, (2005).

[29] H.J. Briegel, D.E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, *Measurement-based quantum computation*, Nature Physics **5**, 19–26 (2009).

[30] D. Hangleiter, M. Kliesch, M. Schwarz, and J. Eisert, *Direct certification of a class of quantum simulations*, Quantum Sci. Technol. **2**, 015004 (2017).

[31] A.R. Meyer and L.J. Stockmeyer, *The Equivalence Problem for Regular Expressions with Squaring Requires Exponential Space*, In Proceedings of the 13th IEEE Symposium on Switching and Automata Theory, pp. 125–129 (1972).

[32] L.J. Stockmeyer, *The polynomial-time hierarchy*, Theor. Comp. Sci. **3**, 1–22 (1977).

[33] A. Ehrenfeucht and M. Karpinski, *The Computational Complexity of (XOR, AND)-Counting Problems*, Technical Report 8543-CS, (1990).

[34] L.J. Stockmeyer, *On Approximation Algorithms for #P*, SIAM J. Comput. **14**, 849–861 (1985).

[35] The proof actually allows for the existence of some multiplicative error, in the form of distributions $Q_{f,\mathrm{Est}}$ which satisfy $|Q_{f,\mathrm{Est}}(\mathbf{s}) - Q_f(\mathbf{s})| < \frac{Q_f(\mathbf{s})}{\mathrm{poly}(n)}$ for all outcomes $\mathbf{s}$, with poly(n) a fixed polynomial. While sampling from such a distribution $Q_{f,\mathrm{Est}}$ is still classically intractable, this is unsatisfactory from a practical standpoint. For example, if any outcome $\mathbf{s}_0$ satisfies $Q_f(\mathbf{s}_0) = 0$, then we must have the probability $Q_{f,\mathrm{Est}}(\mathbf{s}_0)$ be exactly 0 as well. This is clearly impossible to verify for any experimental distribution $Q_{f,\mathrm{Est}}$, leading exact classically intractable sampling results to have a more strained relationship with experimental realities than their average-case counterparts.

[36] D. Gottesman, *The Heisenberg Representation of Quantum Computers*, talk at International Conference on Group Theoretic Methods in Physics (1998), arXiv:quant-ph/9807006.

[37] J. Miller and A. Miyake, *Hierarchy of Universal Entanglement in 2D Measurement-based Quantum Computation*, npj Quantum Information **2**, 16036 (2016).

[38] D. Gottesman and I. L. Chuang, *Demonstrating the Viability of Universal Quantum Computation Using Teleportation and Single-Qubit Operations*, Nature **402**, 390 (1999).

[39] S. Aaronson and L.-J. Chen, *Complexity-Theoretic Foundations of Quantum Supremacy Experiments*, arXiv:1612.05903, (2016).

[40] M. Rossi, M. Huber, D. Bruß, and C. Macchiavello, *Quantum Hypergraph States*, New J. Phys. **15**, 113022 (2013).

[41] O. Gühne, M. Cuquet, F.E.S. Steinhoff, T. Moroder, M. Rossi, D. Bruß, B. Kraus, and C. Macchiavello, *Entanglement and Nonclassical Properties of Hypergraph States*, J. Phys. A **47**, 335303 (2014).

[42] For comparison, the familiar complexity classes P and NP are respectively contained in the zero and first levels of the polynomial hierarchy (PH). While the randomized complexity class BPP has only been proven to lie in the second level of the PH, a proof of the widely conjectured P=BPP would place it in the zero (lowest) level as well. As a corollary, proving P=BPP would allow Stockmeyer counting to be implemented in the second level of the PH, causing the hypothetical collapse invoked in classically intractable sampling results to occur at the second level of the PH, rather than the third.

[43] S. Toda, *PP is as hard as the polynomial-time hierarchy*, SIAM J. Comput. **20**, 865–877 (1991).

[44] On the other hand, the effect of the measurements used in our verification scheme on the measured state is different from that of direct measurements of $h_f^{(i)}$. For example, performing a genuine quantum nondemolition measurement of $h_f^{(i)}$ on the sampling state $|\psi_f\rangle$ would leave it unchanged, whereas our scheme always collapses it to a tensor product of single-qubit X and Z eigenstates. Since we only care about measurement statistics and not the post-measurement state, this has no impact on our protocol.

[45] T. Morimae, Y. Takeuchi, and M. Hayashi, *Verified measurement-based quantum computing with hypergraph states*, arXiv:1701.05688, (2017).

Appendix A: Comparison to Previous Work

We now discuss the relationship of our work to previous proposals for classically intractable sampling with qubits, the class of boson sampling protocols having a largely different flavor with regards to theoretical underpinnings and experimental implementations. As mentioned before, our work is most closely related to that of [8], as it implements their circuit-based IQP sampling in the context of MQC. We have seen that this translation has several practical advantages, mainly that it allows us to use constant depth quantum circuits generated by local interactions to perform classically intractable sampling in constant time. This translation also reveals the role of MQC byproduct operators in simplifying our protocol, with an associated randomness which ends up having no impact on the classical intractability of sampling. Furthermore, the convenient verification scheme utilized in our protocol can be applied equally well in any classically intractable sampling implementation using the IQP sampling states associated with Conjecture 1, revealing an inherent practical advantage of sampling from this class of states. This advantage more generally applies to any protocol which samples from output distributions defined by so-called hypergraph states[40, 41].

Although our work doesn't make use of the alternate Conjecture 2 of [8], concerning the average-case hardness of estimating fully-connected Ising partition functions, our techniques can be easily generalized to define a similar MQC sampling protocol which relies upon Conjecture 2. In this alternate protocol, our CCZ gadget would be replaced by gadgets for the non-Clifford $\sqrt{CZ}$ and T gates, and our byproduct group would contain not only CZ, but also $\sqrt{Z}$ gates. In terms of the Clifford hierarchy of unitary operations [38], the pattern which emerges here is that using gadgets which implement operations at the third level of the Clifford hierarchy leads to a random byproduct group formed from Clifford gates at the second level of the Clifford hierarchy. Just as with our protocol, this would eliminate the need to apply any Clifford gates "by hand", reducing the physical resources needed for sampling.

Our work also has many similarities to the MQC sampling protocol of [13], which similarly runs in constant time using a fixed "brickwork" resource state preparable by a constant depth quantum circuit, and also allows for verification. In our protocol, the average-case hardness of sampling relies on Conjecture 1, while the average-case hardness in [13] relies upon a conjecture regarding the estimation of output probabilities of random quantum circuits, argued in [18] to be a stronger assumption. On the other hand, this latter conjecture is very similar to that used in [12, 19, 39].

While [13] also achieves verification of the hardness of their sampling distribution, their method requires verifying the entire initial MQC resource state. By contrast, our use of Conjecture 1 lets us perform verification in exactly the same manner as sampling, where the only difference is a change in the single-qubit Pauli bases used to perform the final n measurements. This unique duality between sampling and verification arises from the simple byproduct group appearing in our protocol, which is necessary for our preparation measurements to always implement IQP circuits. In contrast, the output states of general random unitary circuits studied in [6, 12] are likely too complicated to allow the associated sampling and verification protocols, or MQC counterparts such as [13, 19], to achieve the duality we observe here.

Appendix B: Randomness of MQC Byproduct Polynomials

Here we study the preparation stage of our MQC protocol, and show that the polynomials $f = \mathbf{a} + \mathbf{b} + \mathbf{c}$ associated with our random output states ρ_f contains uniformly random quadratic and linear coefficients, so that every b_{ij} and c_i is an independent binary random variable with equal $\frac{1}{2}$ probability. We show this by first characterizing the distribution of preparation outcomes $P_{\mathbf{a}}(\mathbf{t})$, where $\mathbf{t} = (t_1, t_2, \ldots, t_m)$, then using this to characterize the distribution $P_{\mathbf{a}}(\mathbf{b} + \mathbf{c})$ of "byproduct polynomials" arising in our protocol. We show that $P_{\mathbf{a}}(\mathbf{b} + \mathbf{c})$ is uniformly random, a fact which holds true in the presence of arbitrary noise with spatial correlations of a bounded distance. This result will be used in our proofs of sampling and verification in Appendices C and D.

We calculate $P_{\mathbf{a}}(\mathbf{t})$ using the Born rule, which in our ideal setting says that given $\mathbf{a}$-dependent preparation measurements on $|\Psi_{\mathrm{Prep}}\rangle$, the probability of obtaining an outcome $|\mathbf{t}_{\mathbf{a}}\rangle$ (where $\mathbf{a}$ denotes the appropriate single-qubit eigenbases) is

$$P_{\mathbf{a}}(\mathbf{t}) = |\langle \mathbf{t}_{\mathbf{a}} | \Psi_{\mathrm{Prep}}\rangle|^2. \tag{B1}$$

The expression $\langle \mathbf{t}_{\mathbf{a}} | \Psi_{\mathrm{Prep}}\rangle$ here denotes not a scalar, but a partial inner product on $|\Psi_{\mathrm{Prep}}\rangle$, consisting of an n-qubit state which isn't measured until the sampling and verification stage of our protocol. Consequently, Eq. (B1) says that $P_{\mathbf{a}}(\mathbf{t})$ is equal to the squared norm of this state $\langle \mathbf{t}_{\mathbf{a}} | \Psi_{\mathrm{Prep}}\rangle$. Although we would expect this output state to be the sampling state $|\psi_f\rangle$, a careful calculation of the inner products arising in our protocol reveals an additional $\frac{1}{\sqrt{2}}$ scalar factor per preparation measurement, as remarked in Figure 1. This shows that $\langle \mathbf{t}_{\mathbf{a}} | \Psi_{\mathrm{Prep}}\rangle = (\frac{1}{\sqrt{2}})^m \, |\psi_f\rangle$, where $f = \mathbf{a} + \mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t})$, which then proves the preparation measurement outcomes to be distributed as $P_{\mathbf{a}}(\mathbf{t}) = 2^{-m}$. We note that this independence of measurement outcomes is a generic feature of MQC state preparation protocols, as the implementation of norm-preserving unitary operations in every preparation measurement will necessarily force Eq. (B1) to take a constant value for all $\mathbf{t}$, corresponding to every preparation outcome t_i being uncorrelated and uniformly random.

We now use the uniform randomness of preparation measurement outcomes $\mathbf{t}$ to prove the uniform randomness of byproduct polynomials $\mathbf{b}+\mathbf{c}$, which depend on $\mathbf{t}$ as $\mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t})$. These global byproducts arise from the local byproduct operators associated with random outcomes t_i in each of the MQC gadgets shown in Figure 1, which are then commuted through our computation to contribute linear and quadratic terms to $\mathbf{b}(\mathbf{t})+\mathbf{c}(\mathbf{t})$. Each quadratic and linear coefficient in $\mathbf{b} + \mathbf{c}$ can thus be expressed as a sum (mod 2) of many different measurement outcomes t_i, and it is clear that the complete randomness of each measurement outcome will lead every byproduct coefficient in $\mathbf{b}+\mathbf{c}$ which contains even a single random t_i to be itself completely random. It is clear that every quadratic coefficient contains contributions from at least one random t_i, with the one exception of b_{1n}. Because our CCZ gadgets only apply CCZ byproduct operators between nearest neighbor logical qubits, and since qubits 1 and n are never adjacent to each other in the circuit diagram of Figure 2, it remains possible that b_{1n} will always be 0. A simple fix for this is to simply vary the ordering among each triple of qubits entering a non-Clifford gadget using $SWAP$ gadgets, so that all qubits are adjacent to all other qubits equally often. In this case, every quadratic coefficient $b_{ij}(\mathbf{t})$ in $\mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t})$ will receive $O(n)$ random contributions from outcomes t_i arising in CCZ gadgets, and every linear coefficient $c_i(\mathbf{t})$ will receive $O(n^3)$ contributions from outcomes arising in 1D cluster wires and $SWAP$ gadgets. This clearly proves that the distribution of byproduct operators will be uniformly random as $P_{\mathbf{a}}(\mathbf{b} + \mathbf{c}) = 2^{-(n_{\mathbf{b}}+n)}$, where $n_{\mathbf{b}} = \binom{n}{2}$.

The above analysis which counts the number of measurement outcomes contributing to each coefficient of $\mathbf{b} + \mathbf{c}$ is unnecessary in an idealized setting, but is useful in the presence of realistic noise and experimental imperfections. We can generally characterize this behavior as a trace preserving quantum operation $\mathcal{E}$ which maps our MQC resource state to some imperfect $\mathcal{E}(|\Psi_{\mathrm{Prep}}\rangle\langle\Psi_{\mathrm{Prep}}|)$. Our measurement statistics $P_{\mathbf{a}}(\mathbf{t})$ in this setting are again set by the Born rule, but now as

$$\begin{aligned} P_{\mathbf{a}}(\mathbf{t}) &= \mathrm{Tr}\left[\mathcal{E}(|\Psi_{\mathrm{Prep}}\rangle\langle\Psi_{\mathrm{Prep}}|)|\mathbf{t}_{\mathbf{a}}\rangle\langle\mathbf{t}_{\mathbf{a}}|\right] && \text{(B2)} \\ &= \mathrm{Tr}\left[|\Psi_{\mathrm{Prep}}\rangle\langle\Psi_{\mathrm{Prep}}|\mathcal{E}^\dagger(|\mathbf{t}_{\mathbf{a}}\rangle\langle\mathbf{t}_{\mathbf{a}}|)\right], && \text{(B3)} \end{aligned}$$

where $\mathcal{E}^\dagger$ represents the quantum operation which is adjoint to $\mathcal{E}$. While $\mathcal{E}^\dagger$ may modify our measurement projectors $|\mathbf{t}_{\mathbf{a}}\rangle\langle\mathbf{t}_{\mathbf{a}}|$ so as to displace or correlate the probabilities of local outcomes t_i, we noted above that the coefficients of byproduct polynomials are determined by at least $O(n)$ different such measurement outcomes, any one of which is capable of completely randomizing the probability of that coefficient. Consequently, in order for noise to alter the distribution of byproduct operators, the operator $\mathcal{E}^\dagger$ must induce correlations between at least $O(n)$ different measurement outcomes in our system. In the presence of any noise with a finite correlation length, this is clearly impossible, which proves the uniform randomness of byproduct operators to be a robust property of our MQC protocol.

Appendix C: Hardness of Approximate Sampling

Here we give a detailed proof of the classical intractability of our MQC sampling protocol under constant variational noise in the output sampling distributions Q_f. We first discuss the general idea behind average-case classically intractable sampling protocols, so as to make clear what precisely needs to be demonstrated in our proof. We then describe the use of classical post-processing on our measurement records to implement "coarse-graining" in the description of our protocol. This coarse-graining lets us simplify the analysis of failure probabilities required in our proof, and eventually lets us prove Theorem 1, with its associated variational error threshold of $\eta_0 = \frac{1}{86}$. We note a certain duality between the proof given here and the proof of Theorem 2 given in Appendix D, with the former using a guaranteed bound on η_0 as a starting point and the latter deriving such a bound on η_0 as an end result.

Any proof of classical intractability of quantum sampling requires adopting somewhat of a dual viewpoint. On the one hand, we recognize that our sampling procedure is an intrinsically quantum task, but at the same time assume that the sampling distributions arising from this quantum process can be exactly replicated using a probabilistic classical algorithm. This assumption, analogous to the assumption of a hidden variable model describing our quantum process, is made in order to derive a (widely conjectured) contradiction, the collapse of the polynomial hierarchy of complexity theory. Even though the probabilities of individual sampling outcomes $Q_f(\mathbf{s})$ are exponentially small and would require exponential time to estimate empirically, if they arise from a classical sampling process, then the technique of Stockmeyer approximate counting can be used to estimate these probabilities up to multiplicative error. In particular, Stockmeyer counting can be used to output an estimate $Q_{f,\mathrm{Est}}(\mathbf{s})$ which is related to our probability of interest by $|Q_{f,\mathrm{Est}}(\mathbf{s}) - Q_f(\mathbf{s})| \leq \frac{Q_f(\mathbf{s})}{\mathrm{poly}(n)}$, for any desired polynomial $\mathrm{poly}(n)$. The use of an average-case complexity conjecture, like Conjecture 1 in our paper, is then required to connect the ability to estimate such probabilities in the presence of noise to the ability to solve #P-hard problems, from which a collapse of the polynomial hierarchy follows.

Stockmeyer counting is an unphysical process which cannot be carried out efficiently using classical or quantum devices, but can be implemented with a hypothetical "alternating Turing machine" capable of efficiently solving problems in the third level of the polynomial hierarchy [42]. Furthermore, Stockmeyer counting involves manipulations on a register of binary random variables underlying our random outcomes, and consequently can only estimate probabilities arising as outcomes of classical randomized algorithms. Nonetheless, if we assume the existence of an efficient classical algorithm for exactly sampling from the distribution $D_f(\mathbf{s}) = \mathrm{ngap}^2(f + \mathbf{s})$, Stockmeyer sampling would then permit a device existing in the third level of the polynomial hierarchy to estimate any $\mathrm{ngap}^2(f)$ up to multiplicative error, and thus solve any problem in #P. Because solving arbitrary problems in #P is known by Toda's theorem [43] to allow one to efficiently solve all problems in the hierarchy, assuming the existence of this efficient classical algorithm for sampling from distributions D_f would necessarily collapse the polynomial hierarchy to its third level, a contradiction. Hence, this proves the task of sampling from arbitrary D_f to be classically intractable.

A necessary ingredient in any *average-case* classically intractable sampling result is a mathematical problem whose estimation remains #P-hard even when our estimates have some finite probability of failing to be multiplicatively close to their actual value. In our setting, this problem is furnished by Conjecture 1, which says that estimating $\mathrm{ngap}^2(f)$ up to $\frac{1}{4}$ multiplicative error is #P-hard, even when a fraction $\epsilon \leq \epsilon_0 = \frac{23}{24}$ of our estimates fail to lie within this $\frac{1}{4}$ multiplicative bound. Evidence in support of Conjecture 1 is given in [8]. This failure probability ϵ_0 ends up determining the allowed deviation of our quantum sampling distributions Q_f from their ideal D_f. If this deviation is sufficiently small, as measured by the variational distance between Q_f and D_f, the assumed computational hardness of estimating $\mathrm{ngap}^2(f)$ then guarantees that our quantum sampling task will be classically intractable. Consequently, our main goal in this proof is to analyze the deviations in our distributions $Q_f(\mathbf{s}) = \mathrm{Tr}(\rho_f |\mathbf{s}_X\rangle\langle\mathbf{s}_X|)$ arising from deviations in our experimental states ρ_f from their ideal $|\psi_f\rangle\langle\psi_f|$, and to find sufficient conditions to guarantee that the failure probability in estimating $\mathrm{ngap}^2(f)$ using Stockmeyer sampling on Q_f is below our threshold ϵ_0.

We now introduce the idea of coarse-grained sampling distributions, which indeed we have already implicitly made use of in the description of our sampling protocol. In Section III, we described different preparation outcomes $\mathbf{t} = (t_1, t_2, \ldots, t_m)$ as giving rise to different ideal sampling states $|\psi_{f(\mathbf{t})}\rangle$ via the correspondence $f(\mathbf{t}) = \mathbf{a} + \mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t})$. This means that whenever different preparation outcomes $\mathbf{t} \neq \mathbf{t}'$ generate the same byproduct polynomials $\mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t}) = \mathbf{b}(\mathbf{t}') + \mathbf{c}(\mathbf{t}')$, the resultant sampling states will be identical. In reality though, it is entirely possible that these preparation outcomes will generate different sampling states $\rho_{\mathbf{a},\mathbf{t}} \neq \rho_{\mathbf{a},\mathbf{t}'}$, leading our description of a single sampling state $\rho_{f(\mathbf{t})}$ to represent a coarse-graining over equivalent preparation outcomes $\mathbf{t}$. In particular, if $P_\mathbf{a}(\mathbf{t})$ denotes the probability of obtaining a preparation outcome $\mathbf{t}$ arising from our $\mathbf{a}$-dependent Pauli measurements on $|\Psi_{\mathrm{Prep}}\rangle$, then we find ρ_f to be given by

$$\rho_f = \frac{1}{P_\mathbf{a}(\mathbf{b}, \mathbf{c})} \sum_{\{\mathbf{t} | \mathbf{b}(\mathbf{t}) + \mathbf{c}(\mathbf{t}) = f + \mathbf{a}\}} P_\mathbf{a}(\mathbf{t}) \rho_{\mathbf{a},\mathbf{t}} \ . \tag{C1}$$

$P_\mathbf{a}(\mathbf{b}, \mathbf{c})$ represents a normalization factor which gives the total probability on input $\mathbf{a}$ of obtaining any outcome $\mathbf{t}$

associated with the byproduct polynomial $\mathbf{b}+\mathbf{c}=f+\mathbf{a}$. While the above coarse-graining might appear trivial, we will now show how this can be used to effectively mix the inequivalent states ρ_f and $\rho_{f'}$ when f and f' differ only in their linear coefficients.

If we describe our overall sampling process at this stage as first preparing a random state ρ_f with $f=\mathbf{a}+\mathbf{b}+\mathbf{c}$, which is then sampled to obtain an X basis outcome of s, then we would record this in an experiment as yielding the outcome $(\mathbf{b},\mathbf{c},\mathbf{s})\in\Omega_{\mathbf{a}}$ in some outcome space $\Omega_{\mathbf{a}}$. From the layout of our sampling protocol, the probability of this outcome is clearly $P_{\mathbf{a}}(\mathbf{b},\mathbf{c},\mathbf{s})=P_{\mathbf{a}}(\mathbf{b},\mathbf{c})Q_f(\mathbf{s})$. Because of the degeneracy $D_{f+\mathbf{s}}(\mathbf{s})=\mathrm{ngap}^2(f)$ for all outcomes $\mathbf{s}$, we say that any such outcome samples from the polynomial f. These exponentially many outcomes are precisely the ones which can be used to obtain an estimate of $\mathrm{ngap}^2(f)$ via Stockmeyer counting, and we will choose our coarse-graining to eliminate this degeneracy, so that each $\mathrm{ngap}^2(f)$ is determined by a unique sampling outcome from a unique output sampling state. We note that this coarse-graining was used implicitly in [8], although interpreted there as an "obfuscation" of output probabilities.

In Appendix B we showed that the distribution of byproduct polynomials is uniformly random as $P_{\mathbf{a}}(\mathbf{b},\mathbf{c})=2^{-(n_{\mathbf{b}}+n)}$, where $n_{\mathbf{b}}=\binom{n}{2}$. Given this robust characterization of $P_{\mathbf{a}}(\mathbf{b},\mathbf{c})$, we will use $\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$ to indicate the conditional probability of obtaining any outcome which samples from $f=\mathbf{a}+\mathbf{b}+\mathbf{c}$, given that the quadratic portion of our byproduct polynomial is $\mathbf{b}$. This leads $\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$ to be

$$\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})=2^{n_{\mathbf{b}}}\sum_{s}P_{\mathbf{a}}(\mathbf{b},\mathbf{c}+\mathbf{s})Q_{f+\mathbf{s}}(\mathbf{s}) \tag{C2}$$

$$=\sum_{\mathbf{s}}2^{-n}\mathrm{Tr}\left(\rho_{f+\mathbf{s}}|\mathbf{s}_X\rangle\langle\mathbf{s}_X|\right) \tag{C3}$$

$$=\mathrm{Tr}\left(2^{-n}\sum_{\mathbf{s}}\rho_{f+\mathbf{s}}|\mathbf{s}_X\rangle\langle\mathbf{s}_X|\right) \tag{C4}$$

$$=\mathrm{Tr}\left(\tilde{\rho}_{\mathbf{a}+\mathbf{b}}|\mathbf{c}_X\rangle\langle\mathbf{c}_X|\right). \tag{C5}$$

We use $|\mathbf{c}_X\rangle$ to indicate the X basis outcome string corresponding to the linear terms of f. In the above, we have also defined $\tilde{\rho}_{\mathbf{a}+\mathbf{b}}$ to be the state

$$\tilde{\rho}_{\mathbf{a}+\mathbf{b}}=2^{-n}\sum_{\mathbf{s}}Z^{\mathbf{s}}\left(\rho_{\mathbf{a}+\mathbf{b}+\mathbf{s}}\right)Z^{\mathbf{s}}, \tag{C6}$$

where $Z^{\mathbf{s}}=\bigotimes_{i=1}^{n}(Z_i)^{s_i}$ indicates a product of Z operators. In the ideal setting where each $\rho_f=|\psi_f\rangle\langle\psi_f|$, the result of applying $Z^{\mathbf{c}}$ to ρ_f is to simply remove the linear components of f, leaving the state $|\psi_{\mathbf{a}+\mathbf{b}}\rangle\langle\psi_{\mathbf{a}+\mathbf{b}}|$. In this idealized setting, the result of averaging over all ρ_f and applying the correction $Z^{\mathbf{c}}$ in each case is to leave the state $\tilde{\rho}_{\mathbf{a}+\mathbf{b}}=|\psi_{\mathbf{a}+\mathbf{b}}\rangle\langle\psi_{\mathbf{a}+\mathbf{b}}|$, which contains only cubic and quadratic terms. While we can't literally implement these unitary corrections $Z^{\mathbf{c}}$ within the setting of MQC, we can simulate their action through classical postprocessing on our measurement outcomes. In particular, whenever we obtain an outcome of $(\mathbf{b},\mathbf{c},\mathbf{s})\in\Omega_{\mathbf{a}}$ in our sampling experiment, we instead record this as a coarse-grained outcome $(\mathbf{b},\mathbf{c}+\mathbf{s})\in\tilde{\Omega}_{\mathbf{a}}$ lying in a simpler outcome space $\tilde{\Omega}_{\mathbf{a}}$. This is equivalent to recording only the polynomial f sampled by our experiment, and forgetting the relative contributions to f from MQC byproduct operators and from sampling outcomes $\mathbf{s}$. The equivalence of this coarse-graining in our measurement records with the action of active unitary corrections arises from the equality $|\mathbf{s}_X\rangle\langle\mathbf{s}_X|=Z^{\mathbf{c}+\mathbf{s}}|\mathbf{c}_X\rangle\langle\mathbf{c}_X|Z^{\mathbf{c}+\mathbf{s}}$ used to derive Eq. (C5).

Given this coarse-grained description of our experiment, we would like to bound the failure probability ϵ of obtaining an estimate $\mathrm{ngap}^2_{\mathrm{Est}}(f)$ which differs from the true $\mathrm{ngap}^2(f)$ by more than a multiplicative factor of $\frac{1}{4}$. By requiring this probability to be less than the $\epsilon_0=\frac{23}{24}$ appearing in Conjecture 1, we will arrive at concrete conditions on our coarse-grained output states $\tilde{\rho}_{\mathbf{a}+\mathbf{b}}$ in order for our MQC protocol to implement classically intractable sampling. While the Stockmeyer counting used to obtain $\mathrm{ngap}^2_{\mathrm{Est}}(f)$ from our sampling probabilities $\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$ technically introduces its own multiplicative error in this estimate, because this error can be reduced in our (hypothetical) Stockmeyer counting algorithm to any inverse polynomial $|\mathrm{ngap}^2_{\mathrm{Est}}(f)-\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})|<\frac{\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})}{\mathrm{poly}(n)}$ while still retaining a polynomial runtime, we will ignore this error in the following and simply set $\mathrm{ngap}^2_{\mathrm{Est}}(f)=\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$.

We first use Markov's inequality to bound the probability of our estimate $\mathrm{ngap}^2_{\mathrm{Est}}(f)$ failing to lie within an arbitrary constant distance of $\mathrm{ngap}^2(f)$, $\mathrm{Pr}_f\left(|\mathrm{ngap}^2_{\mathrm{Est}}(f)-\mathrm{ngap}^2(f)|>2^{-n}\delta\right)$, over arbitrary polynomials $f=\mathbf{a}+\mathbf{b}+\mathbf{c}$. We will later convert this into a failure probability for obtaining an estimate of $\mathrm{ngap}^2(f)$ outside of our allowed $\frac{1}{4}$ multiplicative error. Since the approximate and exact values of $\mathrm{ngap}^2(f)$ can both be interpreted as probabilities in different distributions, $\mathrm{ngap}^2_{\mathrm{Est}}(f)=\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$ and $\mathrm{ngap}^2(f)=D_{\mathbf{a}+\mathbf{b}}(\mathbf{c})$, we find that the distance between these values, averaged over $\mathbf{c}$ with fixed $\mathbf{a}+\mathbf{b}$, is proportional to the variational distance between these distributions as

$$\left\langle|\mathrm{ngap}^2_{\mathrm{Est}}(f)-\mathrm{ngap}^2(f)|\right\rangle_{\mathbf{c}}$$

$$=2^{-n}\sum_{\mathbf{c}}|\tilde{Q}_{\mathbf{a}+\mathbf{b}}(\mathbf{c})-D_{\mathbf{a}+\mathbf{b}}(\mathbf{c})| \tag{C7}$$

$$=2^{-n}\left|\tilde{Q}_{\mathbf{a}+\mathbf{b}}-D_{\mathbf{a}+\mathbf{b}}\right|_1 \tag{C8}$$

Defining $\eta_{\mathbf{a}+\mathbf{b}}=\left|\tilde{Q}_{\mathbf{a}+\mathbf{b}}-D_{\mathbf{a}+\mathbf{b}}\right|_1$ to be the variational distance between these distributions, Markov's inequality then tells us that for any $\delta>0$ and for $f=\mathbf{a}+\mathbf{b}+\mathbf{c}$ with a fixed $\mathbf{a}+\mathbf{b}$,

$$\Pr_{\mathbf{c}}\left(|\mathrm{ngap}^2_{\mathrm{Est}}(f)-\mathrm{ngap}^2(f)|>2^{-n}\delta\right)<\frac{\eta_{\mathbf{a}+\mathbf{b}}}{\delta}, \tag{C9}$$

Having this bound in hand, we now give an anticoncentration bound on the probability that $\frac{1}{4}\mathrm{ngap}^2(f)<2^{-n}\delta$, which lets us convert the above bound into a statement about the failure probability ϵ. We utilize a particular form of Cantelli's inequality stating that for any nonnegative random variable X and constant δ' in $0 \le \delta' \le 1$,

$$\Pr(X \le \delta'\langle X\rangle) \le \frac{\langle X^2\rangle - \langle X\rangle^2}{\langle X^2\rangle - \delta'(2-\delta')\langle X\rangle^2}. \tag{C10}$$

This agrees with the more well-known Paley-Zygmund inequality at $\delta' = 0, 1$, but otherwise gives a more stringent upper bound. Setting $X = \mathrm{ngap}^2(f)$, $\delta' = 4\delta$, and using the result $\langle \mathrm{ngap}^4(\mathbf{a}+\mathbf{b}+\mathbf{c})\rangle_{\mathbf{b},\mathbf{c}} \le 3\cdot 2^{-2n}$ from [8], this lets us restrict the probability of $\frac{1}{4}\mathrm{ngap}^2(f)$ being less than $2^{-n}\delta$ as

$$\Pr_{\mathbf{b},\mathbf{c}}\left(\frac{1}{4}\mathrm{ngap}^2(\mathbf{a}+\mathbf{b}+\mathbf{c}) \le 2^{-n}\delta\right) \le \frac{2}{2+(1-4\delta)^2}. \tag{C11}$$

We now define $\eta = \langle \eta_{\mathbf{a}+\mathbf{b}}\rangle_{\mathbf{a},\mathbf{b}}$ to be the average variational distance between distributions $\tilde{Q}_{\mathbf{a}+\mathbf{b}}$ and $D_{\mathbf{a}+\mathbf{b}}$, averaged over all $\mathbf{a}+\mathbf{b}$. Combining Eq. (C9) with the average of Eq. (C11) over all $\mathbf{a}$, this results in a bound on the multiplicative failure probability of

$$\Pr_f\left(|\mathrm{ngap}^2_{\mathrm{Est}}(f) - \mathrm{ngap}^2(f)| > \frac{1}{4}\mathrm{ngap}^2(f)\right) < \frac{\eta}{\delta} + \frac{2}{2+(1-4\delta)^2}, \tag{C12}$$

which holds for every $0 \le \delta \le \frac{1}{4}$.

We now require the failure probability to be at most $\frac{23}{24}$, in line with Conjecture 1, and numerically optimize over δ to find the largest allowed value of η_0 for which this can be achieved. This yields a maximum of $\eta_0 = 0.01169$, which has a rational lower bound of $\eta_0 \approx \frac{1}{86}$. This completes our proof of Theorem 1.

Appendix D: Verification of Classical Intractability

Here we prove that the verification scheme occurring in the last stage of our MQC protocol does indeed guarantee the classically intractable of our sampling process. We first show that the local X and Z measurements made on our sampling states ρ_f during verification correspond to exact measurements of the nonlocal stabilizers $h_f^{(i)}$, via the parity functions $\pi_f^{(i)}(\mathbf{v})$. This allows us to estimate the average $\langle h_f^{(i)}\rangle_{i,f}$ with respect to random ρ_f, which allows us to bound the average variational distance $\langle |Q_f - D_f|_1\rangle_f$ using results from [30]. If our empirical estimate of $\langle h_f^{(i)}\rangle_{i,f}$ remains sufficiently low, an application of Höffding's inequality lets us show that $O(n^2)$ verification measurements are sufficient to conclude that $\langle |Q_f - D_f|_1\rangle_f \le \frac{1}{86}$ with any fixed statistical significance, proving Theorem 2.

We first briefly review our verification procedure. After preparation of a random ρ_f, we choose with 50% probability to perform either sampling or verification measurements on ρ_f. If verification is chosen, we further choose a random qubit i of ρ_f which is measured in X, while all other $n-1$ qubits are measured in Z. We denote the measurement outcome string by $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, ignoring the fact that v_i is associated with a different measurement basis. We then use our knowledge of the polynomial f associated with ρ_f to compute a parity function of $\mathbf{v}$, $\pi_f^{(i)}(\mathbf{v}) = \partial_i f(\mathbf{v}) + v_i$, where $\partial_i f$ is the polynomial difference $\partial_i f(\mathbf{v}) = f(v_1, \ldots, v_i+1, \ldots, v_n) - f(v_1, \ldots, v_i, \ldots, v_n)$. It is easy to show that $\partial_i f(\mathbf{v})$ is independent of the value of v_i.

We show here that the process of measuring $\mathbf{v}$ using single-qubit Pauli measurements and then computing $\pi_f^{(i)}(\mathbf{v})$ is exactly equivalent to measuring the nonlocal stabilizer $h_f^{(i)}$ as $h_f^{(i)}(\mathbf{v}) = (-1)^{\pi_f^{(i)}(\mathbf{v})}$, where $h_f^{(i)}(\mathbf{v})$ indicates the $h_f^{(i)}$ outcome corresponding to $\mathbf{v}$. Both processes yield binary random variables as their output, and in order to prove that their probability distributions are identical, we can prove that both measurement schemes are associated with identical Hermitian observables. While measurements of $h_f^{(i)}$ are clearly associated with the Hermitian operator $h_f^{(i)}$ itself, it isn't immediately clear how we should interpret the measurements of $\mathbf{v}$ as measuring any particular Hermitian operator. The answer comes by recognizing that our relevant measurement statistics during verification consist only of the binary values $\pi_f^{(i)}(\mathbf{v})$, and forgets the specific outcomes $\mathbf{v}$ which produced them. Translating these $\pi_f^{(i)}$ outcomes into equivalent $h_f^{(i)}$ outcomes shows the expectation value of $h_f^{(i)}$ on ρ_f to be

$$\langle (-1)^{\pi_f^{(i)}(\mathbf{v})}\rangle_{\mathbf{v}} = \sum_{\mathbf{v}\in GF(2)^n} (-1)^{\pi_f^{(i)}(\mathbf{v})} \Pr(\mathbf{v}|\rho_f) \tag{D1}$$

$$= \sum_{\mathbf{v}} (-1)^{\partial_i f(\mathbf{v}) + v_i} \mathrm{Tr}\left[\rho_f \left(H_i|\mathbf{v}\rangle\langle\mathbf{v}|H_i\right)\right] \tag{D2}$$

$$= \mathrm{Tr}\left[\rho_f \left(X_i \sum_{\mathbf{v}} (-1)^{\partial_i f(\mathbf{v})} |\mathbf{v}\rangle\langle\mathbf{v}|\right)\right] \tag{D3}$$

$$= \mathrm{Tr}\left(\rho_f h_f^{(i)}\right). \tag{D4}$$

In the last equality, we have used the definition of $h_f^{(i)}$ in Eq. (3), while in the second to last equality we used $X_i = \sum_{v_i} (-1)^{v_i} H_i |v_i\rangle\langle v_i| H_i$. This reveals that the expectation value of $(-1)^{\pi_f^{(i)}(\mathbf{v})}$ is equal to that of $h_f^{(i)}$ on ρ_f, and since we made no assumptions about ρ_f, this shows that our verification scheme is exactly equivalent to measuring $h_f^{(i)}$ [44].

As a concrete example, suppose we are working with the 3-qubit sampling state $|\psi_{x_1x_2x_3}\rangle = CCZ_{123}|+\rangle^{\otimes 3}$ and wish to measure the stabilizer $h^{(1)}_{x_1x_2x_3} = X_1CZ_{23}$. In this case, we would perform our verification by measuring X on qubit 1, Z on qubits 2 and 3, and then computing the polynomial $\pi_f^{(i)}(\mathbf{v}) = v_1 + v_2v_3$. This process, which can be thought of as obtaining classical values and plugging them in to the stabilizer itself, would indicate a success when $v_1 = 1$ and $v_2 = v_3 = 1$, or when $v_1 = 0$ and at least one of $v_2 = 0$ or $v_3 = 0$ holds true.

Given the ability to measure arbitrary $h_f^{(i)}$ using single-qubit X and Z measurements, we now note that the average $\langle h_f^{(i)}\rangle_i = \frac{1}{n}\sum_i \langle h_f^{(i)}\rangle$ over randomly chosen sites i is equal to 1 on a given ρ_f only when $\rho_f = |\psi_f\rangle\langle\psi_f|$ is the ideal sampling state. More generally, the techniques of [30] show that this average can be used to bound the closeness of ρ_f to $|\psi_f\rangle\langle\psi_f|$, as measured by the fidelity $F_f = \sqrt{\langle\psi_f|\rho_f|\psi_f\rangle}$. For our purposes, it will be more convenient to work with the square of this quantity, F_f^2. When $\langle h_f^{(i)}\rangle_i \geq 1 - \frac{2}{n}$, ρ_f cannot be orthogonal to $|\psi_f\rangle$, and must have a fidelity squared of at least $F_f^2 \geq 1 - \frac{n}{2}(1 - \langle h_f^{(i)}\rangle_i)$. If we average both sides of this equality over polynomials $f = \mathbf{a} + \mathbf{b} + \mathbf{c}$ with random $\mathbf{b}+\mathbf{c}$, then we find that the average fidelity squared $\langle F_f^2\rangle_f$ of output states ρ_f relative to their intended $|\psi_f\rangle$ is bounded by the average $\langle h_f^{(i)}\rangle_{i,f}$ as

$$\langle F_f^2\rangle_f \geq 1 - \frac{n}{2}(1 - \langle h_f^{(i)}\rangle_{i,f}) \tag{D5}$$

With Eq. (D5) in hand, we can now bound the average variational distance $\langle |Q_f - D_f|_1\rangle_f$ between the sampling distributions arising from ρ_f and $|\psi_f\rangle$. We utilize the fact that the quantum 1-norm distance $||\rho_f - |\psi_f\rangle\langle\psi_f|||_1 \geq |Q_f - D_f|_1$ gives an upper bound on the variational distance of any output sampling distributions, where $||\rho_f - |\psi_f\rangle\langle\psi_f|||_1 = \mathrm{Tr}\left(|\rho_f - |\psi_f\rangle\langle\psi_f||\right)$ with $|A|$ the operator absolute value. We also use a well-known bound on the 1-norm distance, $||\rho_f - |\psi_f\rangle\langle\psi_f|||_1 \leq \sqrt{1 - F_f^2}$, which together yield

$$\begin{aligned}\langle |Q_f - D_f|_1\rangle_f &\leq \langle ||\rho_f - |\psi_f\rangle\langle\psi_f|||_1\rangle_f && \text{(D6)}\\ &\leq \left\langle \sqrt{1 - F_f^2}\right\rangle_f && \text{(D7)}\\ &\leq \sqrt{1 - \langle F_f^2\rangle_f} && \text{(D8)}\\ &\leq \sqrt{\frac{n}{2}(1 - \langle h_f^{(i)}\rangle_{i,f})}. && \text{(D9)}\end{aligned}$$

In the above, we used the two bounds mentioned, as well as Jensen's inequality for the concave function $\sqrt{1-X}$ in Eq. (D8). Using the relationship between the average of stabilizers and parity functions, $\langle h_f^{(i)}\rangle_{i,f} = \langle (-1)^{\pi_f^{(i)}(\mathbf{v})}\rangle_{\mathbf{v},i,f} = 1 - 2\langle \pi_f^{(i)}(\mathbf{v})\rangle_{\mathbf{v},i,f}$, this finally lets us show that in order to verify that $\langle |Q_f - D_f|_1\rangle_f \leq \eta_0$, it is sufficient for our parity function average to be below

$$\langle \pi_f^{(i)}(\mathbf{v})\rangle_{\mathbf{v},i,f} \leq \frac{\eta_0^2}{n}. \tag{D10}$$

This gives the bound appearing in Theorem 2.

Although any empirical estimate of $\langle h_f^{(i)}\rangle_{i,f}$ obtained from finitely many measurements of $\pi_f^{(i)}(\mathbf{v})$ isn't guaranteed to accurately reflect its true value, we can bound the closeness of this estimate with high probability using the uniformly random distribution of byproduct operators proved in Appendix B. In particular, this tells us that for any fixed $\mathbf{a}$, the average $\langle h_f^{(i)}\rangle_{i,\mathbf{b},\mathbf{c}}$ over output random byproducts is unbiased towards any fixed ρ_f, and thus is an accurate indicator of the uniform closeness of sampling states. This lets us treat $\langle h_f^{(i)}\rangle_{i,f}$ as a simple binary random variable, and use Höffding's inequality to bound the probability of this estimate deviating too far from the true value of $\langle h_f^{(i)}\rangle_{i,f}$.

Höffding's inequality says that if we obtain an estimate $\tilde{X}$ of a binary random variable X using N independent samples, the probability of the true average $\langle X\rangle$ lying above $\tilde{X}$ by more than ζ is

$$\Pr(\langle X\rangle \geq \tilde{X} + \zeta) \leq \exp(-2\zeta^2 N) \tag{D11}$$

In our case, we choose X to be our random parity function, and ζ to be the difference between our specified tolerance $\frac{(1/86)^2}{n}$, and the more numerically precise tolerance for classically intractable sampling derived in Appendix C, $\frac{(0.01169)^2}{n}$. Setting $N = \mu n^2$, this gives a failure probability of

$$p_F \leq \exp(-(2.9138 \times 10^{-6})\mu^2) = \exp(-O(\mu^2)). \tag{D12}$$

Converting this into a success probability $p = 1 - p_F$ then completes our proof of Theorem 2.

A final remark is given to our means of measuring the highly nonlocal, non-Pauli stabilizers $h_f^{(i)}$ through single-qubit Pauli measurements. This technique can actually be generalized to measure the stabilizers of any sampling state formed by starting with $|+\rangle^{\otimes n}$ and applying an IQP circuit composed of $\sqrt{Z}$, Z, CZ, CCZ, and any higher multiply-controlled Z gates. Since these states include all hypergraph states as special instances, our means of measuring non-Pauli stabilizers can be utilized for the goal of measuring hypergraph stabilizers in [45], as the latter requires the non-Pauli portion of measured stabilizers to have support on a constant number of qubits. Generalizing yet further, we see that the necessary and sufficient condition for a local measurement scheme to exactly replicate measurements of a nonlocal operator M in this manner is that M can be diagonalized in a basis which is a tensor product of single-qubit eigenbases. While this allows us to measure many different multi-qubit operators using only single-qubit measurements, a simple counterexample is given by the Hermitian operator $SWAP$, which cannot be measured in this manner owing to its unique -1 eigenstate being the entangled $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$.