Electronic Colloquium on Computational Complexity, Report No. 144 (2024)

Consumable Data via Quantum Communication

Dar Gilboa*
Google Quantum AI

Siddhartha Jain†
UT Austin

Jarrod McClean
Google Quantum AI

Abstract

Classical data can be copied and re-used for computation, with adverse consequences economically and in terms of data privacy. Motivated by this, we formulate problems in one-way communication complexity where Alice holds some data and Bob holds m inputs, and he wants to compute m instances of a bipartite relation on Alice's data and each of his inputs. We call this the asymmetric direct sum question for one-way communication. We give a number of examples where the quantum communication complexity of such problems scales polynomially with m, while the classical communication complexity depends at most logarithmically on m.

For these examples, data behaves like a consumable resource when the owner stores and transmits it as quantum states. We show an application to a strategic data-selling game, and discuss other potential economic implications.

1 Introduction

As statistical models fitted to large datasets are being usefully applied to problems in various fields of science and engineering [13, 16, 38], the use of proprietary data for training or inference raises concerns of data privacy and adequate compensation for the data owner. The destructive nature of measurement in quantum mechanics has the potential to change this picture. In order to model this scenario we introduce the asymmetric direct sum question in one-way communication complexity. Informally we say a relation $\mathcal{R}$ has an asymmetric direct sum property for communication model $\mathcal{M}$ if the communication complexity of computing $\mathcal{R}(x, y_1) \cdot \mathcal{R}(x, y_2) \cdot \ldots \cdot \mathcal{R}(x, y_m)$ is $\Omega(m)$ times the communication complexity of computing $\mathcal{R}(x, y)$. One would expect examples of this in quantum communication when the state Alice sends to Bob undergoes destructive measurement, and may not be copyable. As such, we refer to problems exhibiting this property as *consumable data problems*[1]. To our knowledge, for communication complexity such a model is studied here for the first time. The work of Hazan & Kushilevitz [26] is superficially similar but the crucial difference is that in our work Alice receives only one input whereas in theirs Alice has m independent instances.

This consumable data property is easily seen to *not* hold for all relations when $\mathcal{M}$ corresponds to deterministic one-way communication. Since Alice's message depends only on x, Bob can copy the message m times and solve each of his instances. On the other hand, we show that for quantum communication there is a scaling of $\widetilde{\Omega}(\sqrt{m})$ for the Hidden Matching problem in Section 5.1. We conjecture this can be improved to $\Omega(m)$ for m not too large. Our results in Sections 5.2 and 5.3 get a scaling of $\widetilde{\Omega}(m)$ for the problems of sampling from a solution of a linear system

*darg@google.com
†sidjain@utexas.edu

[1]To make our definition robust, we precisely define consumable data probltems to be those where the scaling is polyomial in m.

ISSN 1433-8092

and estimating the expectation values of two-outcome observables in a multi-party setting, when the parties holding the observables are restricted to using classical communication and performing single-copy measurements of Alice's messages. For these problems, we also show that classical one-way communication does not exhibit this scaling. Note that the quantum asymmetric direct sum property only holds for relations which have a $\Omega(m)$ separation between quantum and deterministic communication complexity, otherwise the quantum protocol could mimic the deterministic one.

Our results have an interesting economic interpretation, specifically in a setting where a data holder wishes to maximize the payoff of selling data that other parties wish to use for computation, or prevent unauthorized re-use. Models of markets are concerned chiefly with goods that are consumed during the process of economic production, known as *rival* goods. However, for almost a century it has been recognized that data and information also play a vital role in economic processes [47]. The ability to cheaply replicate data has long been recognized as its chief distinguishing characteristics compared to other economic resources, and this *nonrivality* has dramatic consequences [8, 45]. It essentially implies that the (albeit idealized) equilibrium known as perfect competition, in which the price of every good on the market is set by its capacity for increasing output, cannot hold once data is involved. In some sense, one cannot "get their money's worth" when data is traded, unless there is some external enforcement mechanism that sets prices. Such a mechanism may lead to suboptimal resource allocations, and requires trust between the parties involved. Examples of this can be seen in recent proposals for data markets [5, 31]. Nonrivality of data may also disincentivise the creation of novel datasets, which could be of particular concern as the production rate of public high-quality data in certain modalities is far outstripped by the growth rate of training sets for large models [50].

In contrast to the classical picture, the fragile nature of quantum states suggests that classical data encoded in the amplitudes of a quantum state may be destroyed upon use for computation. For this to be the case, one must first show that a problem of interest can be solved with data encoded in this way. In addition, one must argue that the resulting states cannot be replicated in a similar manner to classical data. There is an inherent tension between these two goals, since while no-cloning is trivial for general quantum states [42], this is no longer the case once states are structured. As a simple example, given a computational basis state, it can be measured in the computational basis without disturbing it and subsequently copied, and thus acts analogously to classical data. A less trivial example is that states encoding boolean functions may also be copied in some cases [1]. It is therefore a priori not obvious whether any problems satisfy these competing demands. The nuanced nature of clonability for both the state and the task motivates the formal study of this problem.

For any relation which exhibits a quantum asymmetric direct sum property, the state(s) sent by Alice satisfy both of the requirements outlined above. In Section 6.1 we illustrate the economic consequences of this within the framework of production theory, and show that consumability implies the possibility of perfect competition, which cannot be achieved when unencrypted classical data is used. Additionally, in Section 6.2 we formulate a data market as a strategic game, and show that consumability implies potentially larger payoffs for the data seller.

The problems we consider are based on ones that exhibit exponential quantum communication advantages. One may wonder whether such an advantage immediately implies the asymmetric direct sum property. In Section 5.3 we show that this is not the case, by considering the problem of observable estimation in a two-party setting [32, 43], and using shadow tomography [2].

2 Related Work

2.1 Destructive measurement as a resource

The idea of using uncloneability of quantum states as a feature has a long history, starting with the seminal work of Weisner [51] that introduced the notion of quantum money. However, the states used in construction of quantum money schemes typically do not encode or transmit useful information and can benefit from the computational power of pseudo-randomness in quantum state [30]. While no-cloning is easy to show for states with little or no structure, this notion becomes more subtle for structured states, and in particular ones that might be useful in performing computation. Aaronson considered the question of uncloneablity of states that encode classical boolean functions, a problem known as quantum software copy-protection [1, 3]. He showed that the presence of structure enables such states to be cloned unless computational assumptions are made, and even then cannot be ruled out for states that encode functions that can be efficiently learned. The setting we consider can be seen as a distributed generalization of this problem. In the simplest case, evaluating the function of interest requires not only a quantum state in the possession of one player (or the equivalent classical description), but also an observable in the possession of another player.

2.2 Secure Multi-Party Computation

The ability to prevent the re-use of data for computation can in principle be achieved classically using the tools of secure multi-party computation (MPC). The principal objective of line of work is the evaluation of a function $f(x, y)$ where Alice holds x and Bob holds y, in a manner that ensures the security of each player's input and reveals only $f(x, y)$ to both players. There has been extensive work on this problem in various forms since its formulation by Yao [52, 53] (see [21] for a review). Elegant solutions to this problem are known that involve obfuscating (or garbling) the circuit describing f one gate at a time so as to obscure the inputs of each player [23] (which can be achieved using standard cryptographic primitives such as public-key cryptography) or alternatively based on fully homomorphic encryption [9]. Using MPC, the players can run a protocol that enables the evaluation of $f(x, y)$ for a single pair of inputs. However, if Bob wanted to evaluate $f(x, y')$ for some $y' \neq y$, the validity of any MPC scheme implies that this would generally be impossible without rerunning the protocol. Since this requires the cooperation of both parties, it could allow one party to control the number of times another party can compute functions of their joint inputs.

MPC is incomparable to the consumability of data studied in this work, which relies on the properties of quantum mechanics. MPC has the benefits of being generic and not requiring the constant overheads associated with fault-tolerant quantum computation. However, MPC has a number of drawbacks compared to consumable data, namely (i) it requires multiple rounds of two-way communication [22] whereas our notion of consumable data requires only a single round of one-way communication, (ii) it requires cryptographic assumptions while we give unconditional results and (iii) it requires coordination between parties e.g. in choosing a cryptosystem to use, while our construction requires no such coordination. The best known classical techniques also have overheads associated with them due to the need to encrypt data, which however may not be fundamental.

3 Preliminaries & Notation

We denote by $D^{\rightarrow}$ deterministic classical one-way communication complexity. $R^{\rightarrow}_{\varepsilon}$ denotes randomized one-way classical communication complexity with error probability at most ε, in which players are allowed to share an unlimited number of public random bits that are independent of their inputs. We similarly define by $Q^{\rightarrow}_{\varepsilon}$ one-way quantum communication complexity with error probability at most ε. In all cases the one-way restriction implies that only Alice is allowed to send messages to Bob (if there are multiple Bobs, they can communicate among themselves and we do not consider this as part of the complexity of the problem). When the error is a nonzero constant (say 1/3) we omit the subscript. For formal definitions we refer the reader to textbooks by Nisan & Kushilevitz [35], and Lee & Shraibman [36].

We also consider sampling problems, where the goal is for Bob to produce a sample from a target distribution (or some distribution close to it) given some inputs to Alice and Bob. For this type of problem, we define analogously SR, SQ for the classical (randomized) and quantum communication complexity respectively (with the superscript $\rightarrow$ denoting one-way communication as before).

Definition 3.1 (TV Distance)**.** *The total variation distance between two distributions p, q supported on $\mathcal{X}$ is given by*

$$d_{TV}(p,q) = \sup_{S \subseteq \mathcal{X}} |p(S) - q(S)|$$

When we consider sampling problems, we allow constant error in TV distance between the target distribution and the one sampled by the algorithm. Finally, we denote by A^{+} the pseudoinverse of A.

4 Consumable data

We now define the notion of consumable or rival data. Denote by $\mathcal{X}, \mathcal{Y}, \mathcal{O}$ the space of Alice's inputs, Bob's inputs (or those of a single Bob in case there is more than one), and a space of outputs. Below, we use $P = (\mathcal{R}, \mathcal{P}_P, q)$ to denote a family of relational problems $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{O}$ and a set of protocols $\mathcal{P}_P$. We informally use problem to refer to tuples of this kind. Note that one can construct a similar definition for sampling problems, where for each input the goal is for Bob to output a sample from a specific distribution.

We use $\mathcal{R}^m \subseteq \mathcal{X} \times \mathcal{Y}^m \times \mathcal{O}^m$ to denote the m-Bob relational problem where Alice's input is kept constant but all the m Bobs have distinct inputs. The *goal* is to solve the relation on all of Bob instances with 2/3 probability[2]. Similarly, we use $\mathcal{P}^m_P$ to denote the set of protocols where Alice sends one message and the Bobs are allowed to communicate classically if $q = 0$ and quantumly if $q = 1$. So $P^m = (\mathcal{R}^m, \mathcal{P}^m_P, q)$.

For a problem P, denote $c(P)$ to be the one-way communication complexity of the minimum cost protocol in the set $\mathcal{P}_P$ which solves $\mathcal{R}$. Since we will be modeling scenarios where Alice is selling her data to the Bobs who will be using it for computation, the cost here will be in terms of communication between Alice and the Bobs only.

[2]If $\mathcal{R}$ was a decision problem that could be solved with failure probability δ, one could solve $\mathcal{R}^m$ with failure probability δ as well by simple repetition, incurring a multiplicative overhead logarithmic in m. However, this is no longer the case when considering relations, so this notion of complexity is not finite in general.

Definition 4.1 (Consumable data problem)**.** *A problem P is said to be a consumable data problem if*

$$\frac{c(P^m)}{c(P)} = m^{\Omega(1)}$$

Definition 4.2 (Nonconsumable data problem)**.** *A problem P is said to be a nonconsumable data problem if*

$$\frac{c(P^m)}{c(P)} = m^{o(1)}$$

We refer to the quantity appearing in the lower bound in Definition 4.1 as the *consumability rate* of P. In other words, we can say that a problem P is a consumable data problem if its consumability rate is polynomial, and it is nonconsumable otherwise. There is a subtlety in this definition, in the sense that the benefit of consumability arises when Alice chooses to use a particular communication protocol (typically a quantum one over a classical one) but the definition itself does not specify why she would have such a preference. A natural way to introduce a preference is by formulating a strategic game that involves communication. We provide an example in Section 6.2.

Note that any problem involving only deterministic classical communication must be nonconsumable – every Bob can just copy Alice's message into his own working space. We also show in Section 5.3 that if P corresponds to a *decision* problem, then even with quantum communication it must be a nonconsumable data problem. This is because the Bobs can apply the Shadow Tomography protocol [2] (unless the Bobs are only allowed classical communication between themselves and limited quantum memory). Nevertheless, consumability can be proved for certain search problems (with many solutions) solved using randomised or quantum communication. There are a few cases where consumability or nonconsumability can be characterized, which we discuss below:

Lemma 1. *For any relational problem $\mathcal{R}$ and resource q, if the protocol is deterministic one way classical communication, $P = (\mathcal{R}, D^{\rightarrow}, q)$, then $\frac{c(P^m)}{c(P)} = 1$ and the data is nonconsumable.*

Proof. For the m-Bob problem, Alice sends the *same* message as the protocol for the original problem. Since her message depended only on her input, and must enable Bob to solve the problem for any possible input of his, the message can be re-used m times and the correctness guarantee holds for every instance on Bob's end. □

Similarly,

Lemma 2. *For any relational problem $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{O}$, with $|\mathcal{O}| = K$ and resource q, if the protocol is randomized one-way classical communication, $P = (\mathcal{R}, R^{\rightarrow}, q)$, then $\frac{c(P^m)}{c(P)} = O(K \log m)$.*

Proof. This can be achieved by learning the distribution of Bob's output under the randomness of Alice's message. Using folklore results (see [15]) this can be done using $K \log m$ times the amount of communication, by learning the K-outcome distribution up to error $1/m$. □

Lemma 3. *For any relation with an output space of size K, $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{O}$, with $|\mathcal{O}| = K$, if the protocol is one way quantum communication $P = (\mathcal{R}, Q^{\rightarrow}, q = 1)$ then $\frac{c(P^m)}{c(P)} = \widetilde{O}(K \log^2 m) c(P)$.*

Proof. Akin to Lemma 2, we want to give a protocol for for P^m using a protocol for P as a subroutine. We do this by relying on the work of Gong & Aaronson [24] who proved that the distribution of K-outcome POVMs on $\log N$ qubits can be learned to constant additive error in $\widetilde{O}(K \log^2 m \log N)$ copies. □

All of these lemmas can be generalised to the setting where $\mathcal{P}_P$ is a strict subset of one of these sets, which applies to the sampling models.

We note that we restricted our definition of consumable data to a non-interactive setting of one-way communication. We also focus on unconditional results. If we allow interactivity and computational assumptions, then an analogous capabilities can be (conditionally) achieved by classical communication using Secure Multiparty Communication protocols as outlined in Section 2.2. Our results do not contradict such constructions since in a single application of an MPC protocol, Alice's messages will not allow Bob to solve the problem for any inputs.

5 Examples of consumable quantum data

5.1 Hidden Matching

Hidden matching is a famous example of a relation that exhibits an exponential separation between quantum and classical one-way communication complexity [10]. We prove that for the asymmetric direct sum version of the Hidden Matching problem, quantum data behaves as a consumable resource (while classical data does not).

The original problem is defined as follows:

Problem 1 (Hidden Matching [10])**.** *Alice is given a string $x \in \{0,1\}^N$. Bob is given a perfect matching M over $[N]$. Their goal is for Bob to output $(i, j, x_i \oplus x_j)$ where $(i, j) \in M$. Only Alice is allowed to send messages to Bob.*

One can naturally generalize this problem to the setting of multiple matchings as follows:

Problem 2 (Multiple Hidden Matchings ($\mathrm{MHM}_{N,m}$))**.** *Alice is given a string $x \in \{0,1\}^N$. Each of the m Bobs is given m perfect matchings $\{M_k\}$ over $[N]$. Their goal is to output $(i, j, x_i \oplus x_j)$ where $(i, j) \in M_k$ for all k. Only Alice is allowed to send messages to Bob.*

A tight lower bound shows that classical communication indeed acts like a nonrival good for this problem. While it is known that $R^{\rightarrow}(\mathrm{HM}_N) = \Omega(\sqrt{N})$ [10], we believe this is the first characterization of the deterministic complexity of the Hidden Matching problem. The results are consistent with Lemma 1.

Lemma 4. $D^{\rightarrow}(\mathrm{MHM}_{N,m}) = D^{\rightarrow}(\mathrm{HM}_N) = N/2 + 1$.

Proof: Appendix A.

Now we note that even with randomized communication, this relation does not possess the consumable data property.

Lemma 5. $R^{\rightarrow}(\mathrm{MHM}_{N,m}) = O(\log m) R^{\rightarrow}(\mathrm{HM}_N) = O(\sqrt{N} \log m)$

Proof. We adapt the upper bound for HM_N. Alice sends the values of randomly chosen $O(\sqrt{N} \log m)$ nodes, which by a birthday paradox style calculation and union bound has a constant probability of containing one edge from each matching. □

There is a quantum algorithm that solves this problem with probability 1 using $m \log N$ qubits of communication, which is a trivial repetition of the algorithm of [10]:

Lemma 6. $Q^{\rightarrow}(MHM_{N,m}) = O(m \log N)$.

Proof. Alice sends Bob a copy of the state $|\psi\rangle = N^{-1/2}\sum_{i=1}^{N}(-1)^{x_i}|i\rangle$ over $\log N$ qubits. Denoting the k-th pair in Bob's a matching that Bob holds by (i_k, j_k), Bob measures the state using the N-outcome POVM defined by $E_{k,b} = \frac{1}{2}\left(|i_k\rangle + (-1)^b|j_k\rangle\right)\left(\langle i_k| + (-1)^b\langle j_k|\right)$ for $k \in [N/2], b \in \{0,1\}$. This process is repeated for every matching. □

It is clear that the states in the algorithm above cannot be re-used after a measurement to solve the problem for multiple matchings. Since each POVM has N possible outcomes, approaches based on gentle measurement that are discussed in Section 5.3 should not be applicable to this problem without requiring poly(N) copies of the state.

We also have the following lower bound on the quantum communication required to solve the problem:

Lemma 7. *$Q^{\rightarrow}(MHM_{N,m}) = \Omega(\sqrt{m})$ for $m \leq N/2$.*

Proof. Let us consider the distributional complexity of $\text{MHM}_{N,m}$ where Alice's input is a uniform random string $\mathbf{X} \sim U(\{0,1\}^N)$. The Bobs have a deterministic input Y, where M_1 is just the matching $\{(i, i+1)|i \text{ odd}, i < N\}$. The matching M_k is just the k^{th} cyclic permutation on nodes on the left. The Bobs output random variables $o_k = (i_k, j_k, x_{i_k} \oplus x_{j_k})$ as their respective solutions. For notational convenience, we define $\mathbf{O} = o_1 o_2 \dots o_m$. Note that since $m \leq N/2$, each matching consists of $N/2$ edges that do not appear in any other matching. It follows that for any choice of $\mathbf{O}$, no edge (as defined by the first two entries of each o_k) will be repeated.

Let $\rho_{\mathbf{X}}$ be density matrix corresponding to the message of length l sent by Alice, of dimension 2^l. By Holevo's theorem, $I(\mathbf{X} : \mathbf{O}) \leq l$. We will show that if the Bobs solve $\text{MHM}_{N,m}$ then $I(\mathbf{X} : \mathbf{O}) \geq \Omega(\sqrt{m})$. This gives us the required lower bound.

$I(\mathbf{X} : \mathbf{O}) = H(\mathbf{O}) - H(\mathbf{O}|\mathbf{X})$. Note that $H(\mathbf{O}|\mathbf{X}) = 0$ since every Bob's output is deterministic given the input $\mathbf{X}$. Thus, $I(\mathbf{X} : \mathbf{O}) = H(\mathbf{O})$. To make this tuple amenable to analysis, we remove dependencies in the output by considering a spanning forest of the graph induced by $V = \cup_k\{i_k, j_k\} = \cup_k\{i_k\} \cup \cup_k\{j_k\}$. We have that $|V| \geq \sqrt{m}$ since we had a graph with m distinct edges by construction. Therefore, we get a lower bound of $\Omega(\sqrt{m})$ by Lemma 8. □

Lemma 8. *If we have a tree T on n vertices labelled with variables $x_1 \dots x_n$, then if x is a uniform random string then the set of random variables $P_T = \{b_{uv}|(u,v) \in T\}$ where each $b_{uv} = x_u \oplus x_v$ with probability at least 2/3 has total entropy at least $\Omega(n)$.*

Proof. We prove this by induction on the height of the tree T, say h. If $h = 0$, then we have only 1 vertex and the set of parities is empty so the entropy is 0. In the inductive step, we assume that for all trees of height $h-1$, the statement is true. Now, consider any tree T on n vertices of height h. Let L be the set of leaves of this tree, and set $k = |V(T) \setminus L|$. We know that the subtree of T up to height $h-1$ has total entropy on the set $P_{T,h-1}$ at least $k-1$. For any vertex $v \in L$, let $p(v)$ be its parent. Then since x_v is a uniform random bit and v does not appear in any other parities, $H(x_v \oplus x_{p(v)}|P_{T,h-1}) = 1$. Since $b_{uv} = x_u \oplus x_v$ with probability 2/3, by concavity of entropy we have that $H(b_{uv}|P_{T,h-1}) = \Omega(1)$. We now iterate this argument over all leaves, adding the parities at the leaves to the conditioning. We thus prove our claim. □

Combining Lemma 6 and Lemma 7 gives for $m \leq N/2$

$$\frac{Q^{\rightarrow}(\text{MHM}_{N,m})}{Q^{\rightarrow}(MHM_{N,1})} = \tilde{\Omega}(\sqrt{m}). \tag{5.1}$$

This implies that Multiple Hidden Matchings is a consumable quantum data problem, with consumability rate $\approx \sqrt{m}$. We believe this result can probably be sharpened to $\tilde{\Omega}(m)$ for $m \ll N$.

Note that this is not the case classically. The deterministic lower bound in Lemma 4 also illustrates explicitly that the message sent by Alice to solve a single matching, if composed of raw bits of her input, can be immediately re-used to solve the problem for all possible matchings (since $N+1$ bits will contain the end-points of an edge of any possible perfect matching).

5.2 Linear regression sampling

Another key problem type for which it is possible to transform data into a rival good is sampling problems with a quantum communication advantage. In this type of problem, Alice sends Bob a message, which Bob uses to sample from a target distribution with high accuracy. The essence of the construction is that the quantum communication advantage allows Alice to reveal only a tiny fraction of the original data while allowing Bob to solve the problem, and the method by which he solves it destroys the data that was sent, not allowing it to be reused to generate more samples. We consider here a sampling variant of linear regression introduced by Montanaro et al. [39]:

Problem 3 (Linear Regression Sampling [39] (LRS_N))**.** *Alice is given a vector $x \in \mathbb{S}^{N-1}$. Bob is given a matrix B. The goal is for Bob to produce a sample from the distribution $\mathcal{P}$ over $[N]$ defined by*

$$p_i = \left|\left[B^+ x\right]_i\right|^2 / \left\|B^+ x\right\|_2^2. \tag{5.2}$$

One can naturally generalize this problem to the setting of multiple samples as follows:

Problem 4 (Multiple Linear Regression Sampling ($\mathrm{MLRS}_{N,m}$))**.** *Alice is given a vector $x \in \mathbb{S}^{N-1}$. Bob is given m matrices B_k. The goal is for Bob to produce one sample from each distribution $\mathcal{P}_k$ over $[N]$ defined by*

$$p_i^{(k)} = \left|\left[B_k^+ x\right]_i\right|^2 / \left\|B_k^+ x\right\|_2^2. \tag{5.3}$$

Note that solving the above problem with some inaccuracy η corresponds to sampling from some distribution with total variation error at most η with respect to $\mathcal{P}_k$. In order to consider the communication complexity of these problems, we must first discretize the inputs so that they have finite size. We thus assume all real number are specified to $\log N$ bits of precision. We then have

Lemma 9 ([39])**.** *For constant total variation distance error η in the sampled distribution,*

i) $SR_\eta^{\rightarrow}(\mathrm{MLRS}_{N,1}) = \Omega(N \log N)$.

ii) For any m, $SR_\eta^{\rightarrow}(\mathrm{MLRS}_{N,m}) = O(N \log N)$.

Lemma 10. *i) For TV error $\eta \leq 1/4$, $SQ_\eta^{\rightarrow}(\mathrm{MLRS}_{N,m}) = \Omega(m \log(N/m))$.*

ii) For constant TV error η, $SQ_\eta^{\rightarrow}(\mathrm{MLRS}_{N,m}) = O(m \log(N) \max_k(\left\|B_k^+\right\|^2 / \left\|B_k^+ x\right\|_2^2))$.

Proofs: Appendix A.

While these upper and lower bounds match in terms of their N dependence if $\left\|B_k^+ x\right\|_2$ is relatively large (and in particular does not decay with N), they do not match in terms of their m dependence. One example is when the features of x that different samples are sensitive to are in some sense uniformly distributed, as in the construction used to obtain the lower bound in Lemma 10. In this case, we have $\max_k \left\|B_k^+\right\|^2 / \left\|B_k^+ x\right\|_2^2 = O(m)$. It follows that, restricting to such inputs, we have

$$\frac{SQ_{1/4}^{\rightarrow}(\mathrm{MLRS}_{N,m})}{SQ_{1/4}^{\rightarrow}(\mathrm{MLRS}_{N,1})} = \tilde{\Omega}(m). \tag{5.4}$$

Based on the definitions of Section 4, we obtain that $\mathrm{MLRS}_{N,m}$ is a consumable data problem for quantum data, with consumability rate m.

5.3 Two-outcome observable estimation

In the previous examples, we considered problems that exhibit exponential quantum communication advantages. It is natural to ask if such an advantage implies consumability in some generic sense. We will see that this is not the case when Bob's task is a decision problem.

The examples we discuss here are based on the following problem:

Problem 5 (Vector In Subspace ($\mathrm{VS}_{N,\theta}$) [32])**.** *Alice is given a vector $x \in \mathbb{S}^{N-1}$. Bob is given two orthogonal subspaces of dimension $N/2$ specified by projection operators $M^{(1)}, M^{(2)}$. Under the promise that either $\left\|M^{(1)}x\right\|_2 \geq \sqrt{1-\theta^2}$ or $\left\|M^{(2)}x\right\|_2 \geq \sqrt{1-\theta^2}$ for $\theta < 1/\sqrt{2}$, determine which is the case.*

It is known that Vector in Subspace exhibits an exponential advantage in quantum communication with respect to randomized classical communication complexity [44]. Consider the following generalization:

Problem 6 (Vector In Multiple Subspaces ($\mathrm{VMS}_{N,\theta,m}$))**.** *Alice is given a vector $x \in \mathbb{S}^{N-1}$. Bob is given m pairs of orthogonal subspaces $M_j^{(1)}, M_j^{(2)}$. Given a similar promise to the vector in subspace problem for each pair of subspaces, the goal is to determine which subspaces x has large overlap with.*

The exponential advantage in quantum communication might suggest that for this problem as well, classical data will behave like a nonrival good while the quantum analog might behave like a rival good. This is because even for $m = 1$, Alice must send a significant portion of her input to Bob, and thus she may not be able to derive value that is polynomial in m for larger m. However, the problem can still be solved with relatively little quantum communication, since data states can be re-used in a manner that allows Bob to solve the problem for $m > 1$ with Alice communicating a number of qubits that is only logarithmic in m. This can be achieved via shadow tomography:

Theorem 1 (Shadow Tomography [2] solved with Threshold Search [14])**.** *For an unknown state $|\psi\rangle$ of $\log N$ qubits, given m known two-outcome measurements E_i, there is an explicit algorithm that takes $|\psi\rangle^{\otimes k}$ as input, where $k = \tilde{O}(\log^2 m \log N \log(1/\delta)/\varepsilon^4)$, and produces estimates of $\langle\psi| E_i |\psi\rangle$ for all i up to additive error ε with probability greater than $1-\delta$. $\tilde{O}$ hides subdominant polylog factors.*

$\mathrm{VMS}_{N,\theta,m}$ is a problem of estimating m expectation values up to some constant error (due to the constraint on θ) on a target state. If polynomial error is required, it is known that $\Omega(N)$ qubits of communication may be needed, and hence quantum communication is essentially equivalent to classical communication (from e.g. lower bounds on estimating inner products [18]). Allowing constant error, Theorem 1 says that polylog(m) qubits of communication suffice to solve the problem. This directly implies that, at least if Alice sends multiple copies of her state, a lower bound analogous to Lemma 10 is impossible, as Bob does not require a number of qubits polynomial in m. This shows that an exponential communication advantage is not a sufficient condition for quantum data to behave like a rival good.

Given that multiple entangled copies of a quantum state are known to be a more powerful resource than single copies [27], it would also be interesting to consider a setting where Alice sends only single copies of her data states. One way to do this is by introducing assumptions about

the computation Bob is allowed to perform with his message. It may be possible to remove this assumption by utilizing certified deletion [12]. While requiring additional encryption, this could enable Alice to only send a copy of her state after receiving a certificate that Bob has deleted the previous copy, ensuring that multi-copy measurements cannot be performed.

A key difference between the Vector-in-Subspace problem and the other problems we consider is that the former is a decision problem (a two-outcome measurement), while the latter are sampling problems or relations. This difference was already captured by Lemma 3, where we showed that if the number of outcomes are small then the problem does not exhibit consumability for a large range of parameters. In the next section, we get around this limitation by considering the multiparty setting.

5.4 Multiple Bobs: A communication arms race

The above picture changes when more than two parties are involved. Consider a setting where Alice has a vector which she can encode in a quantum state $|x\rangle$ and each of m Bobs has an observable O_i, Alice is only willing to send the Bobs copies of $|x\rangle$ (when using quantum communication), and the Bobs cannot (i) store multiple copies of $|x\rangle$ or (ii) communicate quantum states between them, this is equivalent to the setting of learning without quantum memory that is studied in [17]. More precisely, this is a setting where each Bob can perform a POVM on a single copy of $|x\rangle$ only, and exchange classical messages which correspond to the classical memory used in this setting. In contrast, the setting of learning with quantum memory (as per [17]) is one where the Bobs are allowed quantum communication (but still can measure only a single copy of $|x\rangle$ each), with the content of the quantum communication channel corresponding to the quantum memory. In both cases, Alice's messages correspond to samples of a quantum state (unknown to Bob) as is standard in learning problems. While the results of [17] apply to samples of a mixed state described by a density matrix ρ, they also apply to a purification of ρ in a larger space. This will not affect the scaling with m which is the main object of interest for our purposes.

Define by $\mathcal{O}$ an ensemble of two-outcome POVMs given by $O_i = U_i Z_n U_i^\dagger$ for $0 \leq i < m/2$ and $O_i = -U_{i-m/2} Z_n U_{i-m/2}^\dagger$ for $m/2 \leq i < m$, where the U_i are drawn i.i.d. from the Haar measure and Z_n acts only on the last qubit.

When only classical communication is used between Alice and the Bobs, an optimal lower bound of $\Omega(\sqrt{N})$ for estimating the expectation value of a single two-outcome observable with constant probability is applicable [25]. Lemma 1 of that paper also provides a matching upper bound in the m-observable case (up to logarithmic factors). Namely, estimating m expectation values of unit norm observables to constant error can be done with probability 2/3 by sending $\tilde{O}(\log(m)\sqrt{N})$ bits from Alice to Bob (where $\tilde{O}$ hides polylog(N) factors). Alice requires no knowledge of the observables themselves. This protocol is based on sending $O(\log(m))$ random stabilizer sketches of Alice's input state $|x\rangle$. Each sketch involves Alice drawing a Clifford unitary C from a uniform distribution over the Clifford group $\mathcal{C}_n$ ($n = \log N$), and computing $\langle 0^{\otimes(n-k)} z|C|x\rangle$ for all $z \in |0,1\rangle^k$ for $2^k = \tilde{O}(\sqrt{N})$. Alice generates $O(\log(m))$ i.i.d. sketches in this way and sends both the measurement results and a description of the Clifford unitaries to the Bobs. Each Clifford unitary is defined by specifying $O(n^2)$ one or two-qubit gates from a small set, and thus has an efficient classical description.

If Alice instead sends copies of her input encoded in the amplitudes of a quantum state $|x\rangle$ to the Bobs, but we allow classical communication only between the Bobs, and restrict the Bobs to performing single-copy measurements, the number of samples of $|x\rangle$ required is linear in m [17]:

Theorem 2 (Corollary 5.7, [17]). *With constant probability over O_i drawn i.i.d. from $\mathcal{O}$, estimating*

	(Qu)bits sent from Alice to the Bobs
Classical A $\to$ Bs, Classical Bs $\leftrightarrow$ Bs	$\tilde{\Theta}(N^{1/2})$ [25]
Quantum A $\to$ Bs, Classical Bs $\leftrightarrow$ Bs	$\tilde{\Theta}(m \log N)$ [28]
Quantum A $\to$ Bs, Quantum Bs $\leftrightarrow$ Bs	$O((\log(m) \log(N))^2)$ [14]

Table 1: A communication arms race in estimating expectation values of two-outcome observables to within constant error: Data behaves as a consumable resource if Alice is only willing to send quantum states encoding her data, while the Bobs can only communicate classically. This ceases to be the case if only classical communication is used, or if the Bobs can communicate quantum states. $\tilde{\Theta}$ hides factors of $\log m$.

the expectation values of all O_i w.r.t. $|x\rangle$ without quantum communication between Bobs with success probability at least $2/3$ requires $\Omega\left(\min\left\{m/\log(m), N\right\}/\varepsilon^2\right)$ copies of $|x\rangle$.

Note that this is worst-case over $|x\rangle$ (if $|x\rangle$ was uniformly random Bobs could just guess 0). Note also that the O_i are chosen so that classical shadows do not help (for the operators in question the Hilbert-Schmidt norm is $||O_i^2|| = N$, which is roughly equivalent to the shadow norm that sets the sample complexity of classical shadows [28]). A matching upper bound (up to $\log(m)$ factors, as long as $m < N$) is obtained by the straightforward approach in which Alice sends each Bob $O(1/\varepsilon_2)$ copies of her state.

When the Bobs are allowed to use quantum communication, we are essentially back to the two-party version of the problem, since they can jointly use shadow tomography [2, 14] to estimate all the expectation values using a logarithmic number of copies of $|x\rangle$. These results are summarized in Table 1.

6 Economic implications of consumable data

6.1 Data as an economic resource in production theory

Production theory [34] is one of the principal frameworks for the quantitative study of economic systems. A fundamental object of interest within this framework is the *production function* $F : \mathbb{R}_+^M \to \mathbb{R}_+$ that quantifies in some form the output of an economic agent, for example the goods produced by a firm. The inputs to F denote the resources required to produce said goods, such as labor, capital and raw materials. For conventional goods of this form, which cannot be replicated at zero cost (and are referred to as *rival* goods), it is known that the production function is typically a degree 1 homogeneous function of its inputs (at least locally when restricted to some set S):

$$F(\lambda x) = \lambda F(x) \tag{6.1}$$

for any $\lambda \geq 0$[3]. This captures the notion that e.g. doubling the number of raw materials will double a firm's output. It follows directly from Euler's theorem for homogeneous functions that within the interior of S,

$$F(x) = x \cdot \frac{\partial F}{\partial x}. \tag{6.2}$$

[3]Strictly speaking, this relationship holds only if each good can serve as a substitute for another, which is a standard assumption.

Since the output of the production function is a measure of the firm's capacity to pay for the needed resources, we see that if the price of resource i, denoted p_i, is set according to

$$p_i = \frac{\partial F}{\partial x_i}, \tag{6.3}$$

for all $i \in [M]$, then the output of the firm suffices exactly to purchase all the resources required, and there is no surplus profit. This is known as competitive equilibrium, which maximizes social welfare in the sense that the price of each good is commensurate to its usefulness in increasing the total output [7, 19].

While it has long been understood at a qualitative level that data is an inherently different resource than the ones considered above due to the ability to copy it for free [47, 8], the quantitative form of this statement was realized decades later by the seminal work of Romer [45]. If we include data y as an input into the production function, we instead have

$$F(\lambda x, y) = \lambda F(x, y) \tag{6.4}$$

rather than the expected need to double each input proportionate to match production as in $F(\lambda x, \lambda y) = \lambda F(x, y)$. This is because the data used by one process can be copied and used by several with negligible additional cost. Euler's theorem once again gives

$$F(x, y) = x \cdot \frac{\partial F}{\partial x}. \tag{6.5}$$

However, since increasing the amount of data will generally increase the output (say by improving the quality of inference), we have $\frac{\partial F}{\partial y} > 0$. It follows that

$$F(x, y) < x \cdot \frac{\partial F}{\partial x} + y \frac{\partial F}{\partial y}. \tag{6.6}$$

Due to this inequality, it is impossible to set prices according to Eq. (6.3). If this were done for all inputs including data, the total output would be insufficient to pay for all the required resources. As a result, markets involving data must be inherently inefficient in the sense that one must underpay for some resource, or must include some external mechanism to enforce adequate compensation for resources that can be freely replicated. Mechanisms such as patent law or subsidies that incentivize innovation are all examples of this. Other examples are afforded by the trusted third parties that are introduced in proposals for data markets and handle the data in lieu of the data buyers themselves [5]. In the context of strategic games that model data selling, the ability to copy data is also manifest in the payoff for the data seller being independent of the number of buyers, unless a mechanism is put in place by which the data buyers all agree to pay in advance for their data [40].

6.1.1 Consumable data as a factor of production

We can interpret the results of Section 5 within this framework (at the limit of large m, N so that m can be considered to be a continuous variable, and computing derivatives with respect to it becomes meaningful). Taking the linear regression sampling problem as an example, the solution of $\mathrm{MLRS}_{N,m}$ is analogous to the output of a production function, with the number of samples m and Alice's message equivalent to λ and y respectively. The result of Lemma 9 is then analogous to Eq. (6.4). Up to constant factors, this is an example of the well-known nonrival nature of classical data. Alice must send a significant portion of her input to Bob for him to produce even a single sample, and once Alice sends her full input he can produce an unlimited number of samples in this

way. If Alice were to sell Bob her data in the setting of a strategic game, her potential payoff will be essentially independent of the value that Bob can derive (since this is proportional to m).

On the other hand, Lemma 10 indicates that if Alice insists on using quantum communication, the data is analogous to a rival good as described by Eq. (6.1). Bob can still produce m samples, but this requires that Alice sends at least a number of qubits proportional to m. If Alice were to charge Bob for each qubit sent for example, she would obtain a payoff proportional to the Bob's output m (as long as $m < N$). The lower bound indicates that this scaling holds regardless of the strategy Alice uses to encode her input into the message, and of the strategy Bob uses to process this message. Using classical resources alone this would be impossible to achieve. We make these notions more precise in the context of a strategic game that models a data market in Section 6.2.

A similar analogy can be made with respect to the Multiple Hidden Matching problem and the multi-party observable estimation problem.

6.2 A posted price data auction with consumable data

We would like to identify more concretely the economic consequences of the consumable nature of quantum data. We consider a formulation naturally related to auction theory [33, 46]. Alice's action space $A_A = \mathbb{R}_+$ is the set of prices she charges for a single bit or qubit of her input. Once Alice fixes a price p, Bob is free to purchase as many bits/qubits as he wants. Bob's action space is thus $A_B = \mathbb{N}$, and we denote the number he purchases by b. This is known as a posted price auction with only a single bidder and multiple items (or a particularly simple combinatorial auction). Assume the number of samples m takes values in $[\overline{m}]$ and Alice has no knowledge of it (say she holds a uniform prior). We also assume the matrices B_i are chosen in a worst-case fashion (in order for our communication lower bounds to be applicable).

For any values of m, p, b, the payoffs of the two players are

$$v_A(m,p,b) = pb, \quad v_B(m,p,b) = \#\mathrm{S}(m,b) - pb, \tag{6.7}$$

where $\#\mathrm{S}(m,b)$ represents the number of samples Bob can produce using a message of b bits/qubits, given that he holds m such B_i).

Consider first the quantum communication case. We know from our lower bound Lemma 10 that for sufficiently large m, there is an absolute constant C such that, if Bob were to purchase b qubits produced by Alice, then

$$\#\mathrm{S}^Q(m,b) \leq \frac{Cb}{\log(N/\#\mathrm{S}^Q(m,b))} \approx \frac{Cb}{\log(N)} \tag{6.8}$$

for some absolute constant C. We also assume $N \gg m$ which allows us to use the approximation $\log N - \log \#\mathrm{S}^Q(m,b) \approx \log N$ since this slightly simplifies the analysis. Since additionally $\#\mathrm{S}(m,b) \leq m$ by definition, we have the upper bound

$$v_B^Q(m,p,b) \leq \min\left\{\frac{Cb}{\log N}, m\right\} - pb. \tag{6.9}$$

If we also assume that Bob's payoff is maximized at the point $b^\star$ that maximizes this upper bound, he is interestsed in solving

$$\max_b \min\left\{\frac{Cb}{\log N}, m\right\} - pb = \begin{cases} m(1 - \frac{p \log N}{C}) & 0 \leq p < \frac{C}{\log N} \quad (b^\star = \frac{m \log N}{C}) \\ 0 & p \geq \frac{C}{\log N} \quad (b^\star = 0) \end{cases} \tag{6.10}$$

with the corresponding value of $b^\star$ in the right column. Alice's payoff is maximized by thus choosing p as close as possible to $C/\log N$ from below without exceeding it, and will be equal to $b^\star(m,p)p = mp\log(N)/C = \tilde{\Omega}(m)$. This holds for any m for which Lemma 10 holds, even though Alice has no knowledge of m.

In the classical case, we know the problem is nonconsumable from Lemma 2. This implies that for $m = 1$, there is a message of length κ independent of m which Alice can send, which Bob can then re-use to produce say ρm samples with some constant probability, for some $\rho \leq 1$.

This implies

$$v_B^C(m,p,b) = \mathbf{1}\left[b \geq \kappa\right]\rho m - pb. \tag{6.11}$$

Bob thus solves

$$\max_b \mathbf{1}\left[b \geq \kappa\right]\rho m - pb = \begin{cases} \rho m - p\kappa & 0 \leq p < \frac{\rho m}{\kappa} \quad (b^\star = \kappa) \\ 0 & p \geq \frac{\rho m}{\kappa} \quad (b^\star = 0) \end{cases} \tag{6.12}$$

Note that unlike the quantum case, Alice has no way of knowing how to choose p appropriately ahead of time, since the critical value below which she receives no payoff depends on m. If she wants to guarantee a nonzero payoff she has to choose $p = \rho/\kappa$ (i.e. assume $m = 1$) in which case her payoff is independent of m.

7 Discussion

We demonstrated that there exist problems for which encoding classical data into quantum states leads to behavior that is akin to that of rival, or consumable, goods, which is generally not possible using classical data alone. The inherent privacy benefits of amplitude-encoded data might also facilitate computation with proprietary data, giving users fine-grained control over the dissemination of their private data without the need for additional encryption. The setup we consider also does not require end-users to possess a quantum computer in order to be valuable. Instead, the user must simply trust an entity possessing a networked quantum computer to distribute data states on their behalf. This is similar to entrusting a bank to distribute funds on the behalf of an account holder. While our results are based on communication complexity, they rely on the properties of the data encoding itself, and thus are also relevant in a scenario where different parties are provided access to the same quantum memory at different times, without requiring networked quantum computers.

Being a preliminary investigation into the possibility of using quantum networks in this manner, our results do not immediately apply to problems with clear economic value. If this were the case, it could enable novel types of data markets and incentive structures for the production of data. It is worth noting however that our results for the linear regression sampling problem apply also to a related problem in which Bob obtains a state that encodes the solution to a linear system rather than a classical sample. Such states are known to be strictly more powerful resources than classical samples [6], and could potentially be useful in learning tasks such as updating the value of a linear estimator with new data (which is typically achieved with the recursive least squares algorithm).

The form of the quantum communication lower bound that indicates the rival behavior of quantum data is reminiscent of a direct sum theorem. Direct sum theorems demonstrate that the complexity of solving m independent instances of certain problems scales linearly with m. They have been studied extensively in both the classical [48, 11, 37] and quantum [49, 29] setting. These results are not directly applicable since in our setting the inputs to Alice are not independent. Thus, this work motivates an *asymmetric* direct sum result for classes of communication relations.

In analogy to the potential clonability of quantum states with structure, there is a sense in which any non-consumable data may be cloned with respect to a particular task sample efficiently,

even when cloning the overall state containing the information remains sample inefficient. This is exemplified by the shadow tomography task above in which the task is solved via the creation of a classical representation of a hypothesis ρ_T, such that $\tilde{r}(E_i\rho_T) \approx \tilde{r}(E_i\rho)$ for all i for the ground truth state ρ. This classical representation ρ_T need not be close in trace distance such that $||\rho - \rho_T||_{\text{tr}}$ is small, as would be required for a high fidelity cloning of the true state. However it suffices for the task of shadow tomography, and admits an entirely classical representation that may be cloned through classical communication at will, making the data non-consumable, hence this task is clonable even when the underlying states might not be.

In restricting access to data that is used for computation, the setting we consider bears some resemblance to that of differential privacy [20, 4]. In differential privacy, a query is promised not to reveal too much about individual datapoints. This is typically achieved classically by adding noise to data, while we achieve a similar capability in spirit by using a noiseless encoding into quantum states.

7.1 Open questions

Our work raises the following open questions in communication complexity:

1. Can the lower bound on the one-way quantum communication complexity of $\text{MHM}_{N,m}$ be improved to $\Omega(m)$ or even $\Omega(m \log N)$ for $m \ll \sqrt{N}$?
2. Can the class of problems with a quantum asymmetric direct sum property be characterized in some generality?
3. Are there explicit problems with asymmetric direct sums for randomized communication? For decision problems, the scaling can be at most $O(\log m)$. Can we get a $\Omega(m)$ scaling for relations?

A remark regarding question 3 is that for a random relation, it is actually impossible to get a constant probability of success for superconstant number of instances because we do not have amplification of success probability. This is why we ask for explicit examples.

Our results are unconditional but restricted to specific problems and one-way communication. By making additional assumptions or utilizing the strategic nature of the problems we consider, it may be possible to extend the class of problems that enable consumable data. Some possible directions are outlined below

Computational assumptions: A setting we have not yet considered, that is touched upon by the task of shadow tomography, is one where the computational power of Bob is restricted. It has been noted that general shadow tomography procedures are expected to scale polynomially with the dimension of the Hilbert space of ρ or the trial state ρ_T. If Bob is restricted to polylog computational time, then the creation of the clonable hypothesis state may become impossible. This is analogous to the effect in cryptographic no-cloning theorems on pseudorandom quantum states [30], where even when sample efficient cloning is possible, no computationally efficient scheme can be used to clone the states of interest. In this context, it is worth noting that learning of certain states that have efficient descriptions, such as pseudorandom states [54], is known to be computationally hard. The addition of computational restrictions on Bob hence potentially widens the class of consumable data tasks, but requires moving beyond a communication complexity model that permits unbounded computation.

Communicating mixed states: Moreover, in the above schemes we have focused on cases where the data is provided as pure states and when the only advantage examined otherwise is communication complexity. When data is transmitted as a pure state, it is known that classical shadows

have the potential to remove strict communication advantages when the circuits are too simple, but reconstruction can still be challenging computationally based on arguments from quantum pseudorandom states [54]. An interesting direction could be to enhance some of the features of the example problems here by providing the data as a mixed state, such that it is more difficult for an adversary to learn under certain assumptions such as lack of a substantial quantum memory. We leave these directions for future work.

Acknowledgements

The authors thank Scott Aaronson and Or Sattath for insightful discussions and comments on the manuscript.

References

[1] Scott Aaronson. Quantum copy-protection and quantum money. In *24th Annual IEEE Conference on Computational Complexity*, pages 229–242. IEEE Computer Soc., Los Alamitos, CA, 2009.

[2] Scott Aaronson. Shadow tomography of quantum states. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 325–338, New York, NY, USA, 2018. ACM.

[3] Scott Aaronson, Jiahui Liu, Qipeng Liu, Mark Zhandry, and Ruizhe Zhang. New approaches for quantum copy-protection. In *Advances in cryptology—CRYPTO 2021. Part I*, volume 12825 of *Lecture Notes in Comput. Sci.*, pages 526–555. Springer, Cham, 2021.

[4] Scott Aaronson and Guy N Rothblum. Gentle measurement of quantum states and differential privacy. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 322–333. ACM, New York, 2019.

[5] Anish Agarwal, Munther Dahleh, and Tuhin Sarkar. A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, New York, NY, USA, 2019. ACM.

[6] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 20–29. ACM, New York, 2003.

[7] Kenneth J Arrow. An extension of the basic theorems of classical welfare economics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, volume 2, pages 507–533, 1951.

[8] Kenneth J Arrow. Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity*, pages 609–626. Princeton University Press, Princeton, 1962.

[9] Gilad Asharov, Abhishek Jain, Adriana López-Alt, Eran Tromer, Vinod Vaikuntanathan, and Daniel Wichs. Multiparty computation with low communication, computation and interaction via threshold FHE. In *Advances in Cryptology – EUROCRYPT 2012*, Lecture notes in computer science, pages 483–501. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[10] Ziv Bar-Yossef, T S Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comput.*, 38(1):366–384, 2008.

[11] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 746–755. IEEE, 2013.

[12] Anne Broadbent and Rabib Islam. Quantum encryption with certified deletion. In *Theory of cryptography. Part III*, volume 12552 of *Lecture Notes in Comput. Sci.*, pages 92–122. Springer, Cham, 2020.

[13] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal,

Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901, 2020.

[14] Costin Bădescu and Ryan O'Donnell. Improved quantum data analysis. In *STOC '21—Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1398–1411. ACM, New York, 2021.

[15] Clément L. Canonne. A short note on learning discrete distributions, 2020.

[16] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, Alex Ray, Raul Puri, Gretchen Krueger, Michael Petrov, Heidy Khlaaf, Girish Sastry, Pamela Mishkin, Brooke Chan, Scott Gray, Nick Ryder, Mikhail Pavlov, Alethea Power, Lukasz Kaiser, Mohammad Bavarian, Clemens Winter, Philippe Tillet, Felipe Petroski Such, Dave Cummings, Matthias Plappert, Fotios Chantzis, Elizabeth Barnes, Ariel Herbert-Voss, William Hebgen Guss, Alex Nichol, Alex Paino, Nikolas Tezak, Jie Tang, Igor Babuschkin, Suchir Balaji, Shantanu Jain, William Saunders, Christopher Hesse, Andrew N Carr, Jan Leike, Josh Achiam, Vedant Misra, Evan Morikawa, Alec Radford, Matthew Knight, Miles Brundage, Mira Murati, Katie Mayer, Peter Welinder, Bob McGrew, Dario Amodei, Sam McCandlish, Ilya Sutskever, and Wojciech Zaremba. Evaluating large language models trained on code. *arXiv [cs.LG]*, 2021.

[17] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. Exponential separations between learning with and without quantum memory. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science—FOCS 2021*, pages 574–585. IEEE Computer Soc., Los Alamitos, CA, 2022.

[18] Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Quantum computing and quantum communications (Palm Springs, CA, 1998)*, volume 1509 of *Lecture Notes in Comput. Sci.*, pages 61–74. Springer, Berlin, 1999.

[19] Gerard Debreu. *Theory of value: An axiomatic analysis of economic equilibrium*, volume 17. Yale University Press, 1959.

[20] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

[21] David Evans, Vladimir Kolesnikov, and Mike Rosulek. A pragmatic introduction to secure multi-party computation. *Found. Trends® Priv. Secur.*, 2(2-3):70–246, 2018.

[22] Rosario Gennaro, Yuval Ishai, Eyal Kushilevitz, and Tal Rabin. The round complexity of verifiable secret sharing and secure multicast. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, New York, NY, USA, 2001. ACM.

[23] O Goldreich, S Micali, and A Wigderson. How to play ANY mental game. In *Proceedings of the nineteenth annual ACM conference on Theory of computing - STOC '87*, New York, New York, USA, 1987. ACM Press.

[24] Weiyuan Gong and Scott Aaronson. Learning distributions over quantum measurement outcomes. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 11598–11613. PMLR, 23–29 Jul 2023.

[25] David Gosset and John Smolin. A compressed classical description of quantum states. In *14th Conference on the Theory of Quantum Computation, Communication and Cryptography*, volume 135 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 8, 9. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019.

[26] Itay Hazan and Eyal Kushilevitz. Two-party direct-sum questions through the lens of multiparty communication complexity. In *31 International Symposium on Distributed Computing*, volume 91 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 26, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017.

[27] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, and Jarrod R McClean. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.

[28] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

[29] Rahul Jain and Srijita Kundu. A direct product theorem for one-way quantum communication. In *36th Computational Complexity Conference*, volume 200 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 27, 28. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2021.

[30] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In *Lecture Notes in Computer Science*, Lecture notes in computer science, pages 126–152. Springer International Publishing, Cham, 2018.

[31] Charles I Jones and Christopher Tonetti. Nonrivalry and the economics of data. *Am. Econ. Rev.*, 110(9):2819–2858, 2020.

[32] Ilan Kremer. *Quantum communication*. PhD thesis, Hebrew University of Jerusalem, 1995.

[33] Vijay Krishna. *Auction Theory*. Academic Press, San Diego, CA, 2 edition, 2009.

[34] Heinz D Kurz and Neri Salvadori. *Theory of Production: A Long-Period Analysis*. Cambridge University Press, 1995.

[35] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1996.

[36] Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends® in Theoretical Computer Science*, 3(4):263–399, 2009.

[37] Troy Lee, Adi Shraibman, and Robert Spalek. A direct product theorem for discrepancy. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 71–80. IEEE, 2008.

[38] Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.

[39] Ashley Montanaro and Changpeng Shao. Quantum communication complexity of linear regression. *ACM Trans. Comput. Theory*, 16(1):Art. 1, 30, 2024.

[40] S Nageeb Ali, Ayal Chen-Zion, and Erik Lillethun. Reselling information. *arXiv [cs.GT]*, 2020.

[41] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)*, pages 369–376. IEEE Computer Soc., Los Alamitos, CA, 1999.

[42] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[43] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of Computing*, STOC '99, pages 358–367, New York, NY, USA, 1999. Association for Computing Machinery.

[44] Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, New York, NY, USA, 2011. ACM.

[45] Paul M Romer. Endogenous technological change. *J. Polit. Econ.*, 98(5):S71–S102, 1990.

[46] Tim Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, Cambridge, England, 2016.

[47] Joseph A Schumpeter. *Capitalism, Socialism, and Democracy*. Harper & Brothers, New York, 1942.

[48] Ronen Shaltiel. Towards proving strong direct product theorems. *Comput. Complex.*, 12(1-2):1–22, 2003.

[49] Alexander A Sherstov. Strong direct product theorems for quantum communication and query complexity. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, New York, NY, USA, 2011. ACM.

[50] Pablo Villalobos, Jaime Sevilla, Lennart Heim, Tamay Besiroglu, Marius Hobbhahn, and Anson Ho. Will we run out of data? an analysis of the limits of scaling datasets in machine learning. *arXiv [cs.LG]*, 2022.

[51] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, 1983.

[52] Andrew C Yao. Protocols for secure computations. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, pages 160–164. IEEE, 1982.

[53] Andrew Chi-Chih Yao. How to generate and exchange secrets. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*, pages 162–167. IEEE, 1986.

[54] Haimeng Zhao, Laura Lewis, Ishaan Kannan, Yihui Quek, Hsin-Yuan Huang, and Matthias C Caro. Learning quantum states and unitaries of bounded gate complexity. *arXiv [quant-ph]*, 2023.

A Omitted Proofs

Proof of Lemma 4. We begin by proving $D^{\rightarrow}(\mathrm{MHM}_{N,1}) \geq N/2+1$.

A deterministic protocol $\mathcal{P}$ for $\mathrm{MHM}_{N,1}$ is defined by a matrix with 2^N rows denoting the inputs to Alice and $(N-1)!!$ columns denoting the inputs to Bob ($(N-1)!!$ is the number of perfect matchings over $[N]$). The entry in the matrix corresponding to inputs (x, M) is a tuple (i, j, b) such that $(i, j) \in M$ and $b = x_i \oplus x_j$. Define by τ a message sent by Alice, and by S_τ the subset of the rows for which Alice sends τ to Bob. The choice of (i, j, b) depends on x only through the message τ. Since the protocol is deterministic, for a given column, the entries in each column of S_τ must have the same value since they share the same τ, M, so we may write (with slight abuse of notation)

$$\mathcal{P}(x, M) = \mathcal{P}(\tau, M) = (i, j, b), \quad (i, j) \in M. \tag{A.1}$$

Thus the rows of S_τ are all identical, and we can view each entry as a constraint that each vector x for which Alice sends the message τ must obey. We will bound the maximal possible size of S_τ by bounding the number of xs that can satisfy all these constraints.

The constraints on the bits can be thought of as edges on a graph $G = (V, E)$ with nodes V indexed by $[N]$. We begin with $E = \emptyset$ and choose a sequence of matchings $\mathcal{M} = \{M^\ell\}$. For every matching, $\mathcal{P}$ must produce a valid output that selects an edge from the matching and constrains the corresponding entries of x. While we have no control over which edge is chosen, we will choose $\mathcal{M}$ in such a way that at each step of the algorithm, the size of the connected components in G increases for any edge output by $\mathcal{P}$.

Denote by $\{C_i^\ell\}$ the connected components of G at step ℓ, and $C^\ell = \bigcup_i C_i^\ell$. Initially we thus have $|C^0| = 0$.

i) $|C^\ell| \leq N/2$

We start with an arbitrary matching M^1. For any x for which Alice communicates τ, the entries in S_τ in the column corresponding to M^1 is S_τ must contain an edge $(i, j) \in M^1$, hence after adding (i, j) to E and M^1 to $\mathcal{M}$ we have $|C^1| = 2$. Denoting by D^ℓ the disconnected nodes, we next define a matching M^2 that pairs each node in C^1 with some node in D^1. The remaining nodes of D^1 are paired among themselves. Note that M^2 cannot be equal to M^1, since M^1 contained an edge between two nodes that are both in C^1 while M^2 does not. We add (i, j) to E where $\mathcal{P}(\tau, M^2) = (i, j, b)$. If the edge connects C^1 and D^1, then $|C^2| = 3$. Otherwise, $|C^2| = 4$.

We pick $M^3, \ldots$ in the same fashion, defining $M^{\ell+1}$ by pairing each node in C^ℓ with a node in D^ℓ (and pairing the remaining nodes arbitrarily). This can be done as long as $|C^\ell| \leq N/2$. At every stage, we are guaranteed that $M^{\ell+1} \notin \mathcal{M}$ by the same argument used for M^2, hence we are assured that it is a valid choice.

After at most $N/2-1$ such steps, we have either $|C^\ell| = N/2+1$ or $|C^\ell| = N/2+2$. From this point a different strategy is required, since there are not enough disconnected nodes in D^ℓ to pair with all the nodes in C^ℓ. Subsequently, we order the nodes in C^ℓ by first ordering the connected components $\{C_i^\ell\}$ by size, with C_0^ℓ being the largest (or tied for the largest, breaking ties arbitrarily), and then arbitrarily ordering the nodes within each C_i^ℓ.

ii) $|C^\ell| > N/2$ and $|C_0^\ell| \leq N/2$

Order the nodes in C^ℓ in the manner specified above. Denote by R_-^ℓ the first $N/2$ nodes in this ordering, and by R_+^ℓ the remaining $|C^\ell| - N/2$ nodes. Define the matching $M^{\ell+1}$ by first pairing each node in R_+^ℓ with a node in R_-^ℓ in descending order (i.e. starting with the nodes in C_0^ℓ). Note that two nodes in the same connected component cannot be paired

in this way. This is because, if this occurred for some connected component C_i, this would imply that either $|C_i| > |C_0^\ell|$ (since C_i must have a node in R_-^ℓ, the boundary between R_-^ℓ and R_+^ℓ divides C_i, so every node in the matching so far is in C_i, and we started the pairing in R_-^ℓ with the nodes in C_0^ℓ and went through all of them and reached C_i) contradicting the imposed ordering, or else $C_i = C_0^\ell$, in which case since some nodes in C_0^ℓ are also in R_+^ℓ, we have $|C_0^\ell| > N/2$ and we terminate the algorithm. Having thus paired all the nodes in R_+^ℓ (we can always do this since $|R_+^\ell| \leq N/2$), we complete $M^{\ell+1}$ by pairing the remaining nodes in R_-^ℓ with the unconnected nodes D^ℓ in an arbitrary way. Note that $M^{\ell+1}$ does not contain any edge between two nodes that are in the same connectivity component. Thus it is distinct from all of the matchings already in $\mathcal{M}$ (since by construction each one contained such an edge) and we can add it to $\mathcal{M}$. We add (j,k) to E where $\mathcal{P}(\tau, M^{\ell+1}) = (j,k,b)$.

For the same reason specified above, the edge from $M^{\ell+1}$ that is selected by $\mathcal{P}$ will either connect two previously unconnected components in C^ℓ hence $C_k^{\ell+1} = C_i^\ell \cup C_j^\ell$ for some i,j,k, or else connect some C_i^ℓ with a previously unconnected edge (meaning $|C^{\ell+1}| = |C^\ell| + 1$).

We run the above algorithm until some step $\tilde{\ell}$ when either (a) $|C^{\tilde{\ell}}| = N$ or (b) $C_0^{\tilde{\ell}} > N/2$.

The algorithm is guaranteed to terminate in $O(N)$ steps. If (a) occurs, then either (a1) there are strictly less than $N/2$ connectivity components or (a2) there are exactly $N/2$ connectivity components, since each one contains at least two nodes. In case (a1), there are strictly less than $N/2$ independent degrees of freedom in the choice of the bits of any x for which Alice sends the message τ, since each connectivity component $C_i^{\tilde{\ell}}$ implies $|C_i^{\tilde{\ell}}|$ constraints of the form $x_j \oplus x_k = b$ where $\mathcal{P}(\tau, M) = (j,k,b), (j,k) \in M$ connects two nodes in $C_i^{\tilde{\ell}}$. In case (a2), there are $N/2$ connectivity components of size 2. We then consider a final matching $M^{\tilde{\ell}+1}$ that first divides $\{C_i^{\tilde{\ell}}\}$ into groups of two $\{K_i\}$ and then pairs each node to a node in a different connectivity component within the same K_i. As before, this matching is valid since $M^{\tilde{\ell}+1} \notin \mathcal{M}$. After including the edge in $\mathcal{P}(\tau, M^{\tilde{\ell}+1})$ into E, G will contain $N/2 - 1$ connected components. As before, there are strictly less than $N/2$ degrees of freedom in choosing x. In case (b), there is a single component of size strictly larger than $N/2$. Thus even if all the remaining nodes are disconnected, there are strictly less than $N/2$ degrees of freedom once again.

In conclusion, in all cases we obtain that the number of rows of S_τ is at most $2^{N/2-1}$. The number of possible messages Alice must send is therefore at least $2^N/2^{N/2-1} = 2^{N/2+1}$ and thus the number of bits Alice must send in order to solve $\mathrm{MHM}_{N,1}$ is at least $N/2+1$. Since this bound is valid for the multi-Bob version of the problem as well, we have $D^{\rightarrow}(\mathrm{MHM}_{N,m}) \geq N/2+1$.

The upper bound is trivial: Alice sends the Bobs the first $N/2+1$ bits of her input. These are sufficient for the Bobs to compute the output for all m matchings simultaneously. The result follows. □

Proof of Lemma 9. i) Theorem 9 of [39], applied to square matrices. The proof is based on lower bounds for distributed Fourier sampling.

ii) It follows from the ability of Alice to send her whole input to Bob to complete the task. □

Proof of Lemma 10. i) Say Alice is given a binary vector y of length $m\log(N/m)$ and there are m Bobs. Each Bob uses the matrix

$$B_j = \sum_{i=(N/m)j}^{(N/m)(j+1)} |i\rangle\langle i| . \tag{A.2}$$

Alice then divides her bits into m sets of size $\log(N/m)$ and treats the bits in each set as an integer $r_j \in [N/m]$. She creates a vector x of length N by concatenating a unary encoding of these numbers, meaning

$$[x_{[(N/m)j:(N/m)(j+1)]}]_i = \sqrt{\frac{1}{m}}\delta_{ir_j}, \tag{A.3}$$

where we used $x_{[l:m]}$ denotes the subset of the entries of a vector ranging from $[l, m)$.

Suppose Alice and the Bobs manage to solve $\mathrm{MLRS}_{N,m}$ with inaccuracy η. This means that Bob produces a sample from a distribution that is at most η in TV from each of his target distributions $\mathcal{P}_j$. From the definition of x and the $B^{(j)}$, $\mathcal{P}_j$ is be a delta function at r_j. This means that with probability at least $1 - 2\eta$, Bob recovers the $\log(N/m)$ bits of r_j by performing a computational basis measurement. It follows that Alice's message to Bob is a random-access encoding of $m \log(N/m)$ bits. From known lower bounds on the number of qubits needed for random access coding [41], if $2\eta < 1/2$, Alice must send at least $\Omega(m \log(N/m))$ qubits to the Bobs.

ii) This follows immediately from the bound of Theorem 4 of [39] with an additional factor of m due to the number of samples, and using $||x||_2 = 1$. The bound uses an amplitude-encoding of x, followed by the application of B_k^+ using block-encoding. If two-way communication is allowed, the complexity can be improved to $O(m \log(N) \max_k \left\|B_k^+\right\| / \left\|B_k^+ x\right\|_2)$ since Alice and Bob can run amplitude amplification.

□

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il