Quantum network utility: A framework for benchmarking quantum networks

Yuan Lee1,2, Wenhan Dai3,4,5, Don Towsley6, and Dirk Englund7,8,9

Edited by Nathan A. Baker, Microsoft Corp, Redmond, WA; received August 15, 2023; accepted February 14, 2024 by Editorial Board Member Peter J. Rossky

The central aim of quantum networks is to facilitate user connectivity via quantum channels, but there is an open need for benchmarking metrics to compare diverse quantum networks. Here, we propose a general framework for quantifying the performance of a quantum network by estimating the value created by connecting users through quantum channels. In this framework, we define the quantum network utility metric \( U_{QN} \) to capture the social and economic value of quantum networks. The proposed framework accommodates a variety of applications from secure communications to distributed sensing. As a case study, we investigate the example of distributed quantum computing in detail. We determine the scaling laws of quantum network utility, which suggest that distributed edge quantum computing has more potential for success than its classical equivalent. We believe the proposed utility-based framework will serve as a foundation for guiding and assessing the development of quantum network technologies and designs.

Quantum networks transmit quantum information between quantum systems separated by large distances, enabling applications like quantum cryptography and quantum sensing that are not possible with classical communication networks alone (1–3). Efforts are underway across the globe to develop the cornerstones of such quantum networks, with the goal of distributing quantum information between quantum memories (4–11). These efforts are diverse, employing a range of protocols and hardware to support unconditionally secure communication, high-accuracy sensing/navigation, and distributed quantum computing. On-demand quantum entanglement is now possible between separated quantum memories (12), and quantum memories have been shown to offer a clear advantage in the quantum secure information capacity compared to direct transmission over equivalent loss channels (13). Recent theory established the secret key capacity for quantum communication networks (14, 15), motivating a diversity of quantum network routing protocols (16–24).

Despite this multitude of applications, a fundamental question remains unanswered: Is there a way to quantify the value of a quantum network? Existing studies that quantify quantum network performance generally focus on quantum communication. For example, ref. 15 established the maximum point-to-point quantum communication capacity. However, it considers channel losses as the only limit on quantum communication rates. In practice, local operation errors and suboptimal link layer network protocols can also limit the performance of quantum networks.

Fundamentally, the value of a network derives from the applications it enables by connecting people and things. A telephone network derives its value from the utility seen by the callers it connects. The value of a datacenter network derives from the added utility of networked, rather than isolated, computers. A wireless sensor network enabled by 5G technology adds value by connecting sensors that, taken in isolation, would be less valuable, i.e., the utility of the whole is greater than the sum of its parts.

Metcalfe’s law states that the value of a classical network increases quadratically in the number of its nodes, or roughly the number of point-to-point connections enabled by the network. However, valuing a quantum network is fundamentally different: For example, connecting multiple quantum computers over a network can yield exponential value with respect to the number of nodes involved when measured by the quantum volume (25, 26). Estimating the value of a quantum network remains an open challenge, due in part to the multiuser applications that are promising for applications like quantum cryptography, distributed computing, and sensing. However, there is no universal method to quantify the value of a quantum network, making it impossible to compare their performance and design more effective networks. Here, we propose a general framework for evaluating quantum networks based on the utility that a network creates for its users. This utility-based approach enables realistic and physically relevant comparisons of different quantum network architectures, protocols, and physical modalities that should inform engineering architectures and help guide policy.

Significance

Quantum networks hold great promise for applications like quantum cryptography, distributed computing, and sensing. However, there is no universal method to quantify the value of a quantum network, making it impossible to compare their performance and design more effective networks. Here, we propose a general framework for evaluating quantum networks based on the utility that a network creates for its users. This utility-based approach enables realistic and physically relevant comparisons of different quantum network architectures, protocols, and physical modalities that should inform engineering architectures and help guide policy.
uniquely enabled by quantum networks. Quantum networks are also able to simultaneously perform different applications, such as distributed computing and sensing, each with complexity that grows exponentially with the number of nodes involved. In the absence of a framework that measures the value of various quantum network-enabled applications, comparing the capabilities of different quantum networks on various tasks remains an unsolved problem. We thus develop a general framework for evaluating quantum networks on a multitude of network tasks performed on two or more nodes.

When it comes to quantifying the worth of a quantum network, we are guided by the same principle as in classical networks. In particular, we consider the value created by the ability to pass quantum information between devices. Specifically, we introduce utility-based metrics to quantify the performance of a quantum network in servicing the diversity of the applications of the envisioned quantum network. In a quantum network, a subset of nodes that perform a task is referred to as a coalition as illustrated in Fig. 1. Each task is associated with a utility function, and the performance of a quantum network is measured by the aggregate utility provided by the network. The utility-based metrics form the general framework for comparing the utility of different quantum networks and can be used to guide the design of quantum networks.

Analyzing these metrics reveals insights in the form of scaling laws for the performance of quantum networks. These laws serve the same purpose for quantum networks as Metcalfe’s law does for classical networks (27): They provide network designers and users with a practical measure of network utility, which in turn informs the expansion of existing networks or the construction of new ones.

After explaining the general framework for quantifying the performance of a quantum network, we provide one example of a quantum network metric that extends IBM’s quantum volume (25) to benchmark quantum computers to quantum networks. This framework can also be used to construct metrics for other common quantum network tasks. By incorporating information on the network’s hardware, connectivity, and link layer protocols, these metrics provide realistic and physically relevant measures of a quantum network’s performance. The metrics also highlight practical considerations for the design of larger quantum networks, so that these networks can generate exponentially higher value for more users at the same time. In fact, we find that distributed quantum computing scales better and is thus likely to be more viable than its classical equivalent.

Our utility-based model for analyzing quantum networks is similar in spirit to the approach that refs. 28 and 29 take to analyze classical networks and that ref. 30 extends to quantum networks. These papers focus on an efficient framework for rate control, whereas we focus on evaluating and comparing networks. However, our analysis extends to the development of rate control algorithms.

1. Results

1.1. Benchmarking Quantum Networks. Quantum networks establish quantum communication channels between distant nodes, which can use them for applications including distributed and/or blind quantum computation, distribution of secure keys, transmission of quantum states, etc. (2, 24, 31–33). For concreteness, we assume that quantum networks establish such channels by distributing entanglement between end users, but our framework also applies to other quantum communication protocols.

A fundamental description of a quantum network’s capabilities is given by its entanglement distribution rate region, which describes the entanglement distribution rates it can simultaneously enable among all groups of users. The rate region captures the scarcity of entanglement resources, as an entangled state that is used to connect one group of users cannot also be used to connect another group of users. The rate region incorporates information on the link-layer and network-layer protocols (17, 18, 24, 34) of the quantum network, such as the routing algorithms between nodes and the error correction procedures at individual nodes.

The quantum communication enabled by the network can be applied to a series of tasks from which users of the network derive value. We attribute a utility function to each of the tasks the quantum network performs. It will be convenient to think of this utility as the users’ willingness to pay for the output of

---

Footnote: *We provide some code to compute the example metric in an interactive notebook, which can be found at https://tinyurl.com/quantumNetworkUtility2022.*
these tasks, but more generally, this abstract “utility” can reflect a monetary value, an equivalent cost of classical cryptography, or an abstract social good. Then, the performance of a quantum network is summarized by the maximum aggregate utility the network can provide.

The rate region alone is poorly suited for comparing the performance of quantum networks because it may not reflect the reality that some communication channels are more valuable than others. Moreover, different quantum networks may connect different groups of users, so their rate regions are incomparable. More importantly, quantum networks may be used for different applications, and the performance of a quantum network must be evaluated in the context of the application for which it is used. The aggregate utility metric incorporates information about how value is derived from the quantum network by assigning utilities to tasks that can be completed using quantum communication. Not only can these utilities guide the design of quantum network protocols but they can also be used to enforce a fair entanglement sharing policy between groups of users.

We introduce some notation for concreteness. We represent the quantum network by the graph \( (\mathcal{V}, \mathcal{E}) \), where \( \mathcal{V} \) is the set of nodes and \( \mathcal{E} \) is the set of physical links connecting pairs of nodes. This entanglement can be used to perform tasks, each of which involves a subset of nodes in \( \mathcal{V} \).

Suppose that the quantum network executes \( D \) tasks repeatedly. The quantum network completes tasks as frequently as the required entanglement distribution rates can be stably supported. Let the feasible task region \( \mathcal{W} \) be the set of task completion rate vectors \( P = (p_1, p_2, \ldots, p_D) \) that can be sustained by the quantum network. Here, \( p_i \) is the rate at which the \( i \)th task is completed. A vector of task completion rates can be sustained if the rate at which the entanglement is consumed lies within the rate region of the network. Users derive value from the quantum network when tasks are completed. In our framework, we allow the utility derived from a task to depend nonlinearly on the rate at which this task is completed, but the utility users enjoy from one task is independent of the rates at which other tasks are completed.

Let \( u_i : \mathbb{R}_+ \to \mathbb{R} \) be the utility function associated with the \( i \)th task. If the quantum network performs task \( i \) at rate \( p_i \), users derive utility \( u_i(p_i) \) from this task. Thus, if the network achieves the task completion rate \( P = (p_1, p_2, \ldots, p_D) \), users derive an aggregate utility \( \sum_{i=1}^{D} u_i(p_i) \) over all tasks.

We choose the task completion rate vector that maximizes the aggregate utility users derive from the network. Note that the feasible task region implicitly depends on the rate region of the network. The quantum network’s performance is measured by the following metric, referred to as the quantum network utility:

\[
U_{\text{QN}} = \max_{P \in \mathcal{W}} \sum_{i=1}^{D} u_i(p_i).
\]  

Different applications of quantum networks call for different tasks and utility functions, thereby giving different quantum network metrics. If a quantum network supports multiple applications, the set of tasks should include the relevant tasks across these applications. Regardless, the metric is defined by an optimization problem over the rate region of the network.

The feasible task region \( \mathcal{W} \) can also account for a rate–fidelity tradeoff through the rate region. SI Appendix, Note 1 provides more details on the connection between the rate region and the feasible task region. Moreover, when the quantum network utility is treated as a dimensioned quantity, it is a universal measure of the value provided by any quantum network. More details can be found in SI Appendix, Note 2.

1.2. Quantum Network Utility for Distributed Quantum Computation. In this section, we present an example of the quantum network utility based on quantum volume proposed in refs. 25 and 35. This utility metric \( U_{\text{comp}}^{\text{QN}} \) quantifies the value derived from performing distributed quantum computing tasks and can be viewed as a “quantum volume throughput.”

1.2.1. Quantum volume. Quantum volume is defined with respect to a computing task known as Heavy Output Generation (HOG) (25). A HOG task comprises some number of layers \( d \). If \( m \) memories are involved in the HOG task, each layer performs \([m/2]\) SU(4) operations over pairs of memories, chosen uniformly at random without replacement. A quantum device is said to perform the \( m \)-memory HOG task up to depth \( d \) if the circuit can be implemented with a given minimum overall accuracy. The value associated with performing an \( m \)-memory, depth-\( d \) HOG task is

\[
u = \beta^{\text{min}(m,d)},
\]

where \( \beta = 2 \) in ref. 25.

Quantum volume is then the maximum value over all HOG tasks that can be performed by the device. Quantum volume can be interpreted as the cost of classical computing needed to simulate the equivalent quantum circuit.

Evaluating the quantum volume of a quantum device is complicated in general, but ref. 25 approximates the quantum volume by stating that an \( m \)-memory, depth-\( d \) HOG task can be performed only if

\[
dm \leq \frac{1}{\epsilon_{\text{eff}}},
\]

where \( \epsilon_{\text{eff}} \) is the average error probability per two-qubit gate.

1.2.2. Extension to networks. Now, we extend quantum volume to a network setting. This quantum network utility metric differs from quantum volume in two main ways.

1. Quantum network utility explicitly considers the rate at which multiqubit operations can be performed because transmitting qubits between nodes incurs significant delays.

2. Quantum network utility accounts for the utility derived simultaneously from tasks performed in different parts of the network, not just the task that generates the highest value.

Without loss of generality, we assume that each node in the network has one single-qubit memory allocated for computation. If a physical node in the network has computation memories consisting of multiple qubits, we can split this node into multiple virtual nodes connected by appropriate local links, such that each virtual node has one computation memory. As before, let \( \mathcal{V} \) be the set of all (possibly virtual) nodes. We also assume that the network only produces bipartite entanglement. This simplifying assumption is realistic for near-term quantum networks.
In the above aggregate utility framework, a task is defined by a coalition of at least two nodes \(\mathcal{M}_i \subseteq \mathcal{V}\) (with \(|\mathcal{M}_i| \geq 2\) performing a HOG task of depth \(d_i\), \(i = 1, 2, \ldots, D\). The utility \(u_i(p_i)\) derived from completing a HOG task \((\mathcal{M}_i, d_i)\) at a given rate \(p_i\) is taken to be proportional to the task completion rate and the quantum volume of the task, i.e.,

\[
u_i(p_i) = \beta^{\min(|\mathcal{M}_i|, d_i)} p_i.
\]

The network utility is then the maximum aggregate utility achieved over all task completion rates in the feasible task region \(\mathcal{W}\):

\[
U^{\text{comp}}_{\text{QN}} = \max_{p \in \mathcal{W}} \sum_{i=1}^{D} u_i(p_i).
\]

Network utility can be interpreted as the quantum volume throughput enabled by the quantum network. Following the interpretation of quantum volume as the equivalent cost of classical computation (25), we can also think of the above metric as the cost savings enabled by the quantum network.

### 1.2.3. Modeling the feasible task region.

We now provide a simple model for the feasible task region, so that we can calculate the network utility of distributed quantum computation. This model maps the rate region to the feasible task region and then constructs the rate region itself.

To construct the rate region, we follow ref. 36 and assume that entanglement swapping occurs with a given efficiency \(q_c\) in node \(c \in \mathcal{V}\). We also assume that, in the absence of any entanglement swapping, the network produces entanglement between nodes \(a\) and \(b\) at rate \(f_{ab}\). Note that if nodes \(a\) and \(b\) are not connected by a physical link, then any communication between these nodes must be generated using entanglement swaps, implying that \(f_{ab} = 0\). Then, the feasible task region can be inserted into the optimization problem for the quantum network utility in Eq. 4. The other variables \(u_{ab}^c\) and \(r_{ab}\) in the description of the feasible task region represent underlying quantities that account for the consumption of entanglement in the network. Intuitively, \(u_{ab}^c\) is the rate at which entanglement between nodes \(a\) and \(c\) is used to generate entanglement between nodes \(a\) and \(b\) via entanglement swaps at node \(c\). Furthermore, \(r_{ab}\) represents the rate of entanglement generated between nodes \(a\) and \(b\), either through a physical link or entanglement swaps.

This optimization problem is a linear program, and the number of variables is polynomial in \(|\mathcal{V}|\) and \(D\). SI Appendix, Note 3 discusses this optimization problem in greater detail.

### 1.2.4. Case study: repeater chains.

Optimization problem Eq. 4 has an issue: The number of possible coalitions, and thus \(D\), grows exponentially with the number of nodes \(|\mathcal{V}|\) in the network. Therefore, the optimization problem may be intractable for a general network. It is desirable to reduce the size of the set of candidate coalitions from which to obtain the utility-maximizing solution.

One method to reduce the size of the candidate set is to determine certain structural properties of the coalitions associated with the optimal solution. This method can be used for general networks, and we next use a repeater chain (shown in Fig. 2) to illustrate it. In repeater chains, the connectivity of physical links corresponds to a chain, so that \(f_{ab} = 0\) for all adjacent \(a, b \in \mathcal{V}\). We say that a coalition is connected if there is a path in the coalition between any two nodes of the coalition. The following proposition shows the connectivity of the coalitions associated with the optimal solution.

**Proposition 1.** There exists an optimal solution to problem Eq. 4 such that any coalition \(\mathcal{M}_i\) with \(p_i > 0\) is connected.

For a repeater chain with \(M = |\mathcal{V}|\) nodes, the number of connected coalitions is \(M(M−1)/2\). This is substantially smaller than the number of coalitions in \(\mathcal{V}\), which grows exponentially with \(M\).

We next investigate the size of the largest coalition. We consider a homogeneous repeater chain, which is a repeater chain with \(f_{ab} = f\) for all adjacent nodes \(a, b \in \mathcal{V}\) and \(q_c = q\) for all nodes \(c \in \mathcal{V}\), where \(f, q > 0\) are constants.

**Proposition 2.** In a homogeneous repeater chain with perfect quantum memories and no gate errors (i.e., \(\epsilon_{\text{eff}} = 0\)), the size of the largest coalition with nonzero task rate in an optimal solution is bounded from below by

\[
M + \log_2 \left( \frac{M^{\log_2 q}}{(1+q)M^3(M−1)^2/4} \right).
\]

Proposition 2 states that the size of the largest coalition increases as \(M − O(\log M)\). Therefore, for large linear chain with perfect memories and gates, almost all nodes in the network should be involved in the same coalition, performing a computation task.

We next consider a case where the quantum network produces entanglement of imperfect fidelity. In this case, the size of the largest coalition depends on the fidelity and does not increase as \(M − O(\log M)\).

**Proposition 3.** In a quantum network with errors (i.e., \(\epsilon_{\text{eff}} > 0\)), the size of the largest coalition with nonzero task rate in an optimal solution is bounded from above by \(1/\sqrt{\epsilon_{\text{eff}}^b}\).

Proposition 3 states that the size of the largest coalition depends on the error rate \(\epsilon_{\text{eff}}\) and not on the size of the network \(M\), assuming the network is sufficiently large. Intuitively, coalitions with size \(|\mathcal{M}_i| > 1/\sqrt{\epsilon_{\text{eff}}^b}\) can only perform HOG tasks up to depth \(d_i \leq 1/(\epsilon_{\text{eff}}|\mathcal{M}_i|) < |\mathcal{M}_i|\). Completing one such task provides the same utility as completing another task with coalition size and depth \(|\mathcal{M}_i| = d_i = d_j < 1/\sqrt{\epsilon_{\text{eff}}^b}\), but consumes more entanglement. Therefore, there is no benefit to execute a HOG task using a coalition with more than \(1/\sqrt{\epsilon_{\text{eff}}^b}\) nodes.

**Proposition 4.** In a homogeneous repeater chain with errors (i.e., \(\epsilon_{\text{eff}} > 0\)), the size of the largest coalition with nonzero task rate in an optimal solution is bounded from below by

\[
m + \log_2 \left( \frac{4^{m \log_2 q} |M/m|}{(1+q)m^3(m−1)(2M−m+1)} \right),
\]

where \(m = \lfloor \sqrt{1/\epsilon_{\text{eff}}^b} \rfloor\).
chains with errors. It states that for sufficiently large $M$ and sufficiently small $\ell$, the size of the largest coalition increases as $1/\sqrt{\epsilon_{\text{eff}}}$ and $O(\log \epsilon_{\text{eff}})$. In other words, the size of the largest coalition stays close to the upper bound of Proposition 3.

Proofs of these propositions are provided in SI Appendix, Note 4.

### 1.3. Numerical Results

In this section, we provide numerical results to demonstrate how we can benchmark quantum networks using the quantum network utility. Two prototypical quantum networks are considered: homogeneous repeater chains and dumbbell networks. Homogeneous repeater chains are defined in the previous section. A dumbbell network (Fig. 3) comprises two star-shaped networks, whose centers are connected by a physical link. We refer to the physical link connecting the centers as the bar and the links connecting the star-shaped networks as spokes.

We allow the base $\beta$ in the quantum volume Eq. 2 to vary to illustrate the effects of varying the utility function. Choosing a larger $\beta$ effectively places a higher value on larger coalitions.

#### 1.3.1. Homogeneous repeater chains.

We analyze the homogeneous repeater chain for two values of $\epsilon_{\text{eff}}$: $\epsilon_{\text{eff}} = 0$ and $\epsilon_{\text{eff}} = 0.1$, corresponding to a realistic error rate. We normalize the rate at which physical links produce entanglement to $f_{i+1} = 0.6$. We take the efficiency of entanglement swapping to be $q = 0.9$.

Fig. 4 shows the quantum network utility, as defined in the optimization problem Eq. 4. Fig. 5 shows the maximum coalition size in the optimal solution to Eq. 4. We observe that with perfect gates ($\epsilon_{\text{eff}} = 0$), the largest coalition includes all nodes in the network. However, with imperfect gates ($\epsilon_{\text{eff}} = 0.01$), the optimal solution does not include coalitions with more than 10 nodes.

These behaviors are consistent with the theoretical results. In general, by Proposition 3, the maximum coalition size is bounded from above by $1/\sqrt{\epsilon_{\text{eff}}}$. This explains the trends in quantum network utility shown in Fig. 4. When $\epsilon_{\text{eff}} = 0$, the largest coalition spans the full chain, and quantum network utility increases exponentially with the length of the chain. However, when $\epsilon_{\text{eff}} > 0$, the largest coalition is bounded above by a constant, so in the limit of long chains, the quantum network utility increases linearly with the length of the chain, like the no-swap value. The scaling behavior of the quantum network volume is similar for different choices of the base $\beta$, pointing to the robustness of the quantum network utility for evaluating the performance of a network.

The entanglement graphs in Fig. 6A and B show the rates at which entanglement between pairs of nodes is consumed in the optimal solution, for repeater chains with $\epsilon_{\text{eff}} = 0$ and $\epsilon_{\text{eff}} = 0.01$ respectively. In Fig. 6A, we observe that if the chain length is sufficiently large, the entanglement graph is fully connected. As entanglement between nodes 0 and $M - 1$ is consumed for a computing task, Proposition 1 implies that there must be a coalition involving all the nodes in the repeater chain. Such behavior is consistent with Proposition 2. Conversely, if the chain is not sufficiently long, only entanglement between adjacent nodes is consumed.

In Fig. 6B, we observe that only entanglement between adjacent nodes is consumed when the chain is sufficiently short.

---

**Fig. 3.** Example of a dumbbell network. The long link connecting the two star-shaped subnetworks is the bar; the other links are the spokes.

**Fig. 4.** Network utility of a homogeneous repeater chain for distributed quantum computation as a function of chain length $M$, for different values of $\beta$. The network utility is scaled to the aggregate utility when no entanglement swaps are performed (i.e., when entanglement from the physical links is used directly). The solid line corresponds to $\epsilon_{\text{eff}} = 0$; the dotted line corresponds to $\epsilon_{\text{eff}} = 0.01$. 

---

**Equation:**

$$U_{\text{QN}} = \max_{(p_i),(r_{ab}), (w_{ab}^m)} \sum_{i=1}^{D} \beta \min(|M_i|, d_i) p_i$$

subject to $p_i \geq 0 \forall i$, $r_{ab} \geq 0 \forall a, b \in V$, $w_{ab}^c \geq 0 \forall a, b, c \in V$;

$$r_{ab} = \sum_{i=1}^{D} 2p_id_i \left\{ \begin{array}{ll}
(\lfloor |M_i| \rfloor - 1)^{-1} & \text{if } |M_i| \text{ is even} \\
|M_i|^{-1} & \text{if } |M_i| \text{ is odd} \\
\end{array} \right\} \mathbb{1}_{a \in M_i} \mathbb{1}_{b \in M_i} \forall a, b \in V;$$

$$|M_i|d_i \leq \frac{1}{\epsilon_{\text{eff}}} \forall i = 1, \ldots, D \text{ such that } p_i > 0;$$

$$r_{ab} \leq f_{ab} + \sum_{c \in V \setminus \{a, b\}} q_c \left( w_{ac}^m + w_{bc}^m \right) - \sum_{c \in V \setminus \{a, b\}} \left( w_{ac}^m + w_{bc}^m \right) \forall a, b \in V;$$

$$w_{ab}^c = w_{ab}^m \forall a, b, c \in V.$$
similarly to Fig. 6A. As chain length increases, larger coalitions are formed. However, unlike Fig. 6A, the largest coalition is limited to 10 nodes. In particular, when $8 \leq M \leq 10$, the entanglement graph is fully connected, but for $M > 10$, the repeater chain is divided into coalitions of up to 10 nodes. This is consistent with Proposition 3, which predicts that gate errors will limit the maximum coalition size to $\left\lfloor \frac{1}{\sqrt{\epsilon_{\text{eff}}}} \right\rfloor = 10$.

1.3.2. Dumbbell networks. We now consider a dumbbell network and focus on the entanglement distribution rate of the bar link relative to the spoke links. We fix the rate at which physical spoke links produce entanglement to be $f = 0.6$ in arbitrary units, and we vary the ratio of the bar rate to the spoke rate. The dumbbell network is assumed to be balanced, with $M_{\text{side}}$ spoke nodes on each side of the bar. The total number of nodes in the network is $M = 2M_{\text{side}} + 2$.

Fig. 7 shows the quantum network utility of the dumbbell network as a function of the bar rate. As the bar rate increases, we find that network utility initially increases faster than the aggregate utility if no swaps are performed, but the network utility eventually increases more slowly than the no-swap utility. Initially, as the bar rate increases, the network can perform more computation tasks involving the full network, but when entanglement from the spokes is exhausted, additional entanglement along the bar link can only contribute to computations involving the two bar nodes. This behavior is reflected in the entanglement graphs in Fig. 8. In the limit where the bar rate dominates the spoke rate, the ratio of the network utility to the no-swap utility converges to 1.

We also find that, for a sufficiently small network, quantum network utility under imperfect gates ($\epsilon_{\text{eff}} = 0.01$) is the same as that under perfect gates ($\epsilon_{\text{eff}} = 0$). In such networks, the coalition that contains all network nodes does not exceed the upper bound on coalition sizes $|M_i| \leq 1/\sqrt{\epsilon_{\text{eff}}}$. However, for large $M_{\text{size}}$, having imperfect gates restricts the size of the largest coalition that can be formed, thus significantly reducing quantum network utility. Such behavior mirrors that of the repeater chain (Fig. 4).

1.4. Scaling Laws of Multiuser Applications. We now analyze the impact of multiuser applications like distributed quantum computation on the value of quantum networks. To do so, we obtain scaling laws for the quantum network utility of
distributed quantum computation using the analysis of the preceding sections. These scaling laws not only allow us to determine the conditions under which larger quantum networks can generate exponentially more value but they also form a stark comparison with classical networks.

1.4.1. Error requirements of scaling quantum networks. We start by focusing on the case of quantum repeater chains, for which we have analytical results. As before, let \( M \) denote the size of the network and \( m = \left\lfloor \sqrt{M/\epsilon_{\text{eff}}} \right\rfloor \) be the upper bound on coalition sizes due to two-qubit gate errors. We also denote by \( M^* \) the size of the largest coalition with nonzero task rate. Propositions 2 and 4 state that \( M^* \) is close to \( \min(M, m) \).

Most of the entanglement generated by the network will be allocated to the largest coalition as it produces the most utility per entanglement consumed. Therefore, the largest coalition also makes the largest contribution to the overall quantum network utility. As \( M \) increases, the quantum network utility of the repeater chain scales as

\[
U_{QN}^{\text{comp}} \sim \beta^M \left( \frac{M}{M^*} \right). 
\]

When \( \epsilon_{\text{eff}} \) is a constant independent of \( M \), quantum network utility stops growing exponentially with \( M \) when \( M = m \), switching instead to linear growth when \( M > m \).

Suppose instead that \( \epsilon_{\text{eff}} \) decreases as \( M \) increases. For quantum network utility to grow exponentially with \( M \) indefinitely, it suffices for \( \epsilon_{\text{eff}} \) to decay as \( O(1/M^2) \). If \( \epsilon_{\text{eff}} \) decreases more quickly, the scaling of quantum network utility is unaffected, but if \( \epsilon_{\text{eff}} \) decreases more slowly, quantum network utility will grow subexponentially with \( M \) instead.

In summary, as quantum networks grow in size, two-qubit gates must improve in fidelity for quantum networks to grow exponentially in utility.

1.4.2. Viability of distributed quantum computing at scale. For general quantum networks, we find that a coalition of size \(|\mathcal{M}_i|\) generates a utility of \( \sim \beta^{|\mathcal{M}_i|/2} \) from at most \(|\mathcal{M}_i|^2\) entanglements. The utility per entanglement increases exponentially with the size of the coalition \(|\mathcal{M}_i|\), indicating that distributed quantum computation uses resources more efficiently when tackling larger computations.

2. Discussion

The framework we have presented above can also be used to answer questions about resource allocation and the commercial viability of quantum networks. It also raises many questions, such as how the utility derived by users of quantum network services can actually be measured, and what market structures could emerge in the quantum network sector as it matures. We briefly address these questions below.

2.1. Utility Function Selection. The coalition results we derived for repeater chains and dumbbell networks are justified within the context of using the quantum volume as a utility function. Beyond distributed quantum computing, other quantum network applications could call for different utility functions. In turn, different utility functions could lead to different optimal coalitions and resource allocations. For example, the utility function for quantum key distribution could be related to the value of distributing secure keys vs. the cost of alternative classical cryptography schemes. For distributed quantum sensing, the utility could depend on the precision improvement over classical sensing.

The network benchmarking framework is designed so that these alternative utility functions can easily be incorporated. If a task \( i \) in the task completion rate vector \( P \) has a utility \( u_i \) different from the quantum volume utility, it suffices to simply replace it in Eq. 1. Therefore, the choice of quantum volume as the utility function should be seen as an initial case study. Exploring alternative utility functions is an important direction for future research.

2.2. Resource Allocation. It is clear that, to maximize the value derived from a quantum network, we should allocate communication resources so as to achieve the optimal task completion rate vector \( P^* = (P_i^*|_{i=1}^D = \arg \max_{P \in \mathcal{W}} \sum_{i=1}^D u_i(p_i)) \). The value derived from a quantum network is then the quantum network utility \( U_{QN} = \sum_{i=1}^D u_i(p_i^*) \). One major reason to build a quantum network is to exploit the quantum communication it enables. To determine whether we should build a quantum network for this reason, we compare the quantum network utility \( U_{QN} \) to the total cost of building and operating the quantum network. If \( U_{QN} \) exceeds the cost of the quantum network, we should allocate resources to its construction.
Of course, this analysis ignores other costs like research and development, which can be significant before quantum networks reach a high technology readiness level. We also assume that the variable costs associated with operating a quantum network are negligible, so that the cost of achieving the task completion rate vector $P^\text{opt}$ does not depend on the task completion rates themselves.

We are also interested in the allocation of resources between users of the quantum network. A common interpretation of the aggregate utility function $\sum_{i=1}^{D} u_i(p_i)$ is the willingness to pay of the users who value task completion the most. The allocation of resources between users can be achieved by the following procedure: Complete the $i$th task for users who enjoy utility exceeding $u_i(p_i^*)/p_i^*$ per unit task completion rate.

2.3. Commercial Viability of a Quantum Network. It would be reasonable to leave the provision of quantum network services to the free market when quantum network technologies approach maturity and we seek to exploit quantum communication for commercial applications. A similar scenario has previously been studied in the context of nonnetworked quantum computing (37), in which quantum computing services compete with classical computing services. In contrast with prior work, quantum network services often have no classical equivalent. Taking this into account, we propose a simple model for the private provision of quantum network services. As before, we assume that the operation of a quantum network incurs negligible variable costs.

We consider a market with a monopolistic network operator. This operator is the sole seller of quantum network services. We assume the network operator has no ability to price discriminate.

One option is for the network operator to sell services in the form of elementary entanglement. Users then use this entanglement to perform the tasks from which they derive value. In the simplest pricing scheme, the network operator charges a monthly price of $x$ per unit entanglement generation rate. Then, a user will purchase entanglement from the network operator if the utility they derive from a task exceeds $x$ times the total amount of entanglement needed for the completion of the task. The network operator will choose to supply entanglement at the price $x$ which maximizes its revenue, and thus profit.

Another option is for the network operator to sell task completions to users instead. We can treat each task as a separate market, and the network operator seeks to maximize its total revenue, and thus profit, across all markets. If the network operator supplies completions of the $i$th task at a price of $x_i$ per unit task completion rate, then users only purchase completion of the $i$th task if they derive a utility of more than $x_i$ per unit task completion rate.

Operating the quantum network is commercially viable if total revenue exceeds total costs. Total revenue depends on the shape of the utility functions $u_i$. Note that, in general, total revenue is lower than the quantum network utility $U_{QN}$. If the cost of providing quantum network services exceeds the total revenue but is lower than the quantum network utility, then operating the quantum network is socially desirable but not commercially viable. In this case, public investment in quantum network services would be essential.

Even if the quantum network is commercially viable, private provision may still not be allocatively efficient. For general utility functions $u_i(p_i)$, neither of the monopolistic markets necessarily achieves the optimal allocation.

Nevertheless, we can show that if the utility functions $u_i(p_i)$ are linear in $p_i$, i.e., if $u_i(p_i) = \alpha_i p_i$ for some $\alpha_i > 0$, then a market in which a profit-maximizing, monopolistic network operator sells task completions achieves the optimal resource allocation $p_i^*$. However, a market where entanglement is the quantum network service to be sold may not achieve the optimal resource allocation. We would need $\alpha_i$ to also be proportional to the amount of entanglement consumed per task completion for the entanglement market to achieve the optimal allocation.

2.4. Measuring Utility. Our framework invites an economic interpretation of the quantum network utility as the social benefit derived from the use of the quantum network. To accurately measure this social benefit, we would need to obtain empirically meaningful measures of the utility users derive from quantum network services.

Theoretically motivated prescriptions of the utility functions $u_i(p_i)$ may not be realistic in practice. For example, in Section 1, we used the quantum volume $v = \beta^{\text{min}(|M|,d)}$ as a measure of a computation task’s utility. An argument could be made that this prescription is not wholly appropriate, as it implies that the utility of a computation task increases exponentially with the number of memories involved, assuming the error rate is sufficiently low. This exponential scaling is unsustainable: The users of a quantum network may be willing to pay twice as much to perform distributed quantum computing over three nodes instead of two, but they may not be willing to pay twice as much to compute over twenty-one nodes instead of twenty. A possible fix is to taper off large quantum volumes above some $v_0$ in the expression for the quantum network utility, using the modified volume $v/(1 + v/v_0)$ in place of $v = \beta^{\text{min}(|M|,d)}$ in the optimization problem Eq. 4. However, this fix does not resolve fundamental issues with utility functions that have not been empirically validated.

Moreover, quantum network utility should also take into account other network-enabled applications, such as agreement protocols (38), distributed sensing (39), and other known (2, 3) and unknown applications, but it is unclear what utility functions are appreciated for these tasks.

To resolve these issues, we propose two distinct but nonexclusive principles for measuring utility in practice.

The first principle is to consider the cost of alternatives to using a quantum network. For example, instead of communicating using quantum-network-distributed secret keys, one could employ quantum-safe classical cryptography, or purchase insurance and accept the risk of eavesdropper attacks. The prices of these alternatives are relatively more well-established than the value of quantum key distribution.

The second principle is to estimate the demand curve, which describes the relationship between prices and the quantity of quantum network services demanded. Demand curves allow us to recover the utility derived from quantum network services. Even though there are many practical approaches to demand estimation (40, 41), it is likely to be difficult in the absence of established markets for quantum network services. However, demand estimation will be more viable as quantum networks become more mature.

2.5. Multidimensional Performance Characteristics. Quantum networks, like classical networks, will generally have performance characteristics that are multidimensional in nature, including factors such as latency, throughput, fidelity, and

---

† This result would apply to the utilities described in Section 1.
connectivity. Despite the advantages of quantum network utility, there could be some scenarios in which a single-letter utility may not adequately capture trade-offs between these different characteristics.

- Context-specific Applications: Depending on the application, whether it be distributed quantum computing, quantum key distribution (QKD), or quantum sensing, the utility of the network may need to account for different aspects of performance that are critical to that application. For instance, in QKD, security and key rate may be more important than the raw entanglement rate that a single-letter utility might emphasize. For quantum communication networks, an obvious 2-D metric may be based on both fidelity and rate.
- Resource Allocation: The utility of a quantum network can also depend on the allocation of limited resources. For example, the optimal distribution of entanglement across different tasks can be a complex problem that is not easily reduced to a single utility figure, especially when considering dynamic network conditions and varying task priorities.
- Scalability Concerns: As quantum networks scale up, a single-letter utility may not accurately reflect the complexity of larger networks. The performance of large-scale networks is influenced by numerous factors, including topology, error rates, and maintenance costs, which may require a composite utility measure.
- Economic and Societal Impacts: The utility of a quantum network might also be evaluated in terms of its broader economic and societal impacts. These are multifaceted and may involve considerations such as the cost–benefit ratio of deploying quantum versus classical technologies, the impact on privacy and security, and the enablement of different types of services and applications. Therefore, it is important to also consider the development of a more comprehensive set of metrics for evaluating quantum network performance. Developing a framework that integrates multiple utility measures into an overarching evaluation scheme could provide a more nuanced understanding of a quantum network’s capabilities and value.

2.6. Market Structure. When quantum network services are provided by private operators, the number of operators and the nature of their competition, i.e., the market structure, determine the resource allocation achieved by the quantum network. Some market structures allocate resources efficiently between competing tasks and users, whereas others could be allocatively inefficient or commercially unviable. The legal framework supporting the private or public provision of quantum network services could also affect market outcomes, as it did in classical telecommunications networks (42).

Even though a legal and economic analysis of different market structures is beyond the scope of this paper, we believe that the development of an ecosystem to allocate quantum communication resources efficiently will be beneficial.

3. Conclusion

We have presented a general framework for benchmarking quantum networks, starting from the rate region as a description of a network’s capabilities. The resulting aggregate utility metric not only takes realistic errors into account but also facilitates comparison across quantum networks with different nodes and/or technologies. The aggregate utility metric can also be interpreted as the social value provided by the quantum network to its users.

We have also developed an example of an aggregate utility metric for distributed quantum computing. This metric extends the quantum volume from quantum devices to quantum networks. We develop some theoretical and numerical results for the quantum network utility in prototypical quantum networks and illustrate its scaling behavior with and without gate errors.

The detailed examples we developed account for the value users derive from quantum communication through distributed quantum computing. Our framework is designed to incorporate further applications of quantum communication, such as key distribution and sensing, simply by specifying the utility associated with completing these other tasks. The framework can also be used to guide the design of quantum networks, by choosing the hardware and protocols that maximize the completion of tasks users value most.

We believe that the adoption of the quantum network utility framework and related aggregate utility metrics will facilitate forecasts of the value of quantum networks at different levels of maturity and help design quantum networks to maximize their near-term and long-term impacts.

Data, Materials, and Software Availability. All study data are included in the article and/or SI Appendix. The code used for simulations and figures in this paper can also be found at https://github.com/leeuyan13/NetworkBenchmarking (43).

Acknowledgments. We thank Gerald Adams (SpaceX), Professor Saikat Guha (University of Arizona), Professor Liang Jiang (University of Chicago), Dr. Jay M. Gambetta (IBM), and Professor Francesco Bova (University of Toronto) for helpful discussions. This work was supported in part by the ARO MURI W911NF2110325, the NSF Center for Quantum Networks, the NSF grant CNS-1955834, and the NSF to the Computing Research Association for the CI Fellows 2020 Program.

Author affiliations: aDepartment of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139; bResearch Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139; cCollege of Information and Computer Sciences, University of Massachusetts, Amherst, MA 01003; and dQuantum Photonics Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139.

1. J. I. Cirac, P. Zoller, H. J. Kimble, H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network. Phys. Rev. Lett. 78, 3221–3224 (1997).
2. S. Wehner, D. Elkouss, R. Hanson, Quantum internet: A vision for the road ahead. Science 362 (2018).
3. W. Sasaki, T. Yamamoto, K. Tsuda, Quantum key distribution: Basic concepts, apparatus, and recent progress. Rev. Mod. Phys. 81, 1365–1444 (2009).
4. D. Bouwmeester et al., Experimental quantum teleportation. Nature 390, 575–579 (1997).
5. J. Laurat, K. S. Choi, H. Deng, C. W. Chou, H. J. Kimble, Heralded entanglement between atomic ensembles: Preparation, decoherence, and scaling. Phys. Rev. Lett. 99, 180504 (2007).
6. M. Sasaki et al., Field test of quantum key distribution in the tokyo QKD network. Opt. Express 19, 10387 (2011).
7. J. Lee, E. Besins, A. Dahlberg, S. Wehner, D. Englund, A quantum router architecture for high-fidelity entanglement flows in quantum networks. npj Quant. Inf. 8, 75 (2023).
8. T. Y. Chen et al., Implementation of a 46-node-quantum metropolitan area network. npj Quant. Inf. 7 (2021).
9. S. L. N. Hermans et al., Qubit teleportation between non-neighbouring nodes in a quantum network. Nature 605, 663–668 (2022).
10. L. M. Duan, M. D. Lukin, J. I. Cirac, P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics. Nature 414, 413–418 (2001).
11. N. Sangouard et al., Long-distance entanglement distribution with single-photon sources. Phys. Rev. A 76, 050301 (2007).
12. P. C. Humphreys et al., Deterministic delivery of remote entanglement on a quantum network. Nature 558, 268–273 (2018).
13. M. K. Bhaskar et al., Experimental demonstration of memory-enhanced quantum communication. Nature 580, 60–64 (2020).
14. S. Pirandola, R. Laurenza, C. Ottaviani, L. Banchi, Fundamental limits of repeaterless quantum communications. Nat. Commun. 8, 15043 (2017).
15. S. Pirandola, End-to-end capacities of a quantum communication network. Commun. Phys. 2, 51 (2019).
16. S. Muralidharan et al., Optimal architectures for long distance quantum communication. Sci. Rep. 6, 20463 (2016).
17. M. Pant et al., Routing entanglement in the quantum internet. npj Quant. Inf. 5, 25 (2019).
18. A. Dahlberg et al., “A link layer protocol for quantum networks’’ in Proceedings of the ACM Special Interest Group on Data Communication, SIGCOMM 2019 (Association for Computing Machinery, New York, NY, USA, 2019), pp. 159–173.
19. B. C. Coutinho, W. J. Munro, K. Nemoto, Y. Omar, Robustness of noisy quantum networks. Commun. Phys. 5 (2022).
20. L. Bugalho, B. C. Coutinho, F. A. Monteiro, Y. Omar, Distributing Multipartite Entanglement over Noisy Quantum Networks. Quantum 7, 920 (2023).
21. V. Dobrovitski, G. Fuchs, A. Falk, C. Santori, D. Awschalom, Quantum control over single spins in diamond. Annu. Rev. Condens. Matter Phys. 4, 23–50 (2013).
22. F. Flamini, N. Spagnolo, F. Sciarrino, Photonic quantum information processing: A review. Rep. Progr. Phys. 82, 016001 (2018).
23. M. A. Nielsen, I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, UK, 2000).
24. H. Choi, M. G. Davis, Á. G. Iñesta, D. R. Englund, Scalable quantum networks: Congestion-free hierarchical entanglement routing with error correction (2023).
25. A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, J. M. Gambetta, Validating quantum computers using randomized model circuits. Phys. Rev. A 100 (2019).
26. Y. Zhang, D. Niu, A. Shabani, H. Shapourian, Quantum volume for photonic quantum processors. Phys. Rev. Lett. 130 (2023).
27. B. Metcalfe, Metcalfe’s law after 40 years of ethernet. Computer 46, 26–31 (2013).
28. F. Kelly, Charging and rate control for elastic traffic. Eur. Trans. Telecommun. 8, 33–37 (1997).
29. F. P. Kelly, A. K. Maulloo, D. K. H. Tan, Rate control for communication networks: Shadow prices, proportional fairness and stability. J. Operat. Res. Soc. 49, 237–252 (1998).
30. G. Verdoyan, S. Wehner, Quantum Network Utility Maximization. arXiv [Preprint] (2022). https://arxiv.org/abs/2210.08135 (Accessed 29 October 2022).
31. J. I. Cirac, P. Zoller, H. J. Kimble, H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network. Phys. Rev. Lett. 78, 3221–3224 (1997).
32. H. J. Briegel, W. Dur, J. I. Cirac, P. Zoller, Quantum repeaters: The role of imperfect local operations in quantum communication. Phys. Rev. Lett. 81, 5932–5935 (1998).
33. M. Rul, N. H. Han, H. Chui, D. Englund, R. Hanson, Quantum networks based on color centers in diamond. J. Appl. Phys. 130, 070901 (2021).
34. A. Pirker, W. Dur, A quantum network stack and protocols for reliable entanglement-based networks. New J. Phys. 21, 033003 (2019).
35. H. K. C. Beukers et al., Qurebb (2023). https://github.com/QuTech-Delft/QuREBB.
36. W. Dai, T. Peng, M. Win, Optimal remote entanglement distribution. IEEE J. Selected Areas Commun. 38, 540–556 (2020).
37. F. Bova, A. Goldfarb, R. G. Melko, Quantum economic advantage. Manage. Sci. 69, 1116–1126 (2023).
38. M. Epping et al., Multi-partite entanglement can speed up quantum key distribution in networks. New J. Phys. 19, 093012 (2017).
39. E. Khribouline, J. Borregaard, K. D. Greve, M. Lukin, Optical interferometry with quantum networks. Phys. Rev. Lett. 123 (2019).
40. J. Hausman, G. Leonard, J. D. Zona, Competitive analysis with differentiated products. Ann. d’Economie et de Stat., 159–180 (1994).
41. A. Nevo, Measuring market power in the ready-to-eat cereal industry. Econometria 69, 307–342 (2001).
42. J. A. Hausman, W. E. Taylor, Telecommunication in the US: From regulation to competition (almost). Rev. Indus. Org. 42, 263–280 (2013).
43. Y. Lee, W. Dai, Analysis code. Github. https://github.com/leeyuan13/NetworkBenchmarking. Accessed 7 April 2024.