Draft genome sequence of Nitrosomonas sp. ANs5, an extremely alkali-tolerant ammonia-oxidizing bacterium isolated from Mongolian soda lakes

Valera, Jose R.; Sorokin, Dimitry Y.; Santoro, Alyson E.

DOI
10.1128/mra.00330-25

Publication date
2025

Document Version
Final published version

Published in
Microbiology Resource Announcements

Entanglement improves coordination in distributed systems

Francisco Ferreira da Silva
Delft Networks and QuTech, Delft University of Technology
Delft, The Netherlands
francisco@delftnetworks.com

Stephanie Wehner
QuTech, Delft University of Technology
Delft, The Netherlands
s.d.c.wehner@tudelft.nl

ABSTRACT

Coordination in distributed systems is often hampered by communication latency, which degrades performance. Quantum entanglement enables correlations stronger than classically possible without communication. Such correlations manifest instantaneously upon measurement, irrespective of the physical distance separating the systems. We investigate the application of shared entanglement to a dual-objective optimization problem in a distributed system comprising two servers. The servers process both a continuously available, preemptible baseline task and incoming paired customer requests, to maximize the baseline task throughput subject to a Quality of Service (QoS) constraint on average customer waiting time. We present a rigorous analytical model demonstrating that an entanglement-assisted routing strategy allows the system to achieve higher baseline throughput compared to communication-free classical strategies, provided the baseline task's output exhibits sufficiently increasing returns with processing time. This advantage stems from entanglement enabling better coordination, which allows the system to satisfy the customer QoS constraint with a lower overall probability of splitting customer requests, leading to more favorable conditions for baseline task processing and thus higher throughput. We further show that the magnitude of this throughput gain is particularly pronounced for tasks exhibiting increasing returns, where output grows super-linearly with processing time. Our results identify optimization of scheduling in distributed systems as a novel application domain for near-term quantum networks.

QuNet '25, September 8–11, 2025, Coimbra, Portugal

ACM ISBN 979-8-4007-2097-0/25/09.
https://doi.org/10.1145/3749096.3750029

CCS CONCEPTS

• **Networks** → **Packet scheduling**; • **Hardware** → **Quantum communication and cryptography**;

KEYWORDS

quantum networks, entanglement-assisted coordination, distributed systems, scheduling algorithms, non-local games

ACM Reference Format:
Francisco Ferreira da Silva and Stephanie Wehner. 2025. Entanglement improves coordination in distributed systems. In *Quantum Networks and Distributed Quantum Computing (QuNet '25), September 8–11, 2025, Coimbra, Portugal.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3749096.3750029

1 INTRODUCTION

Coordination enables efficient operation in distributed systems. An important application lies in scheduling, where incoming requests must be assigned across multiple servers [4]. Optimal load balancing often relies on global state information, such as current server loads. However, acquiring this information via classical communication introduces latency. In latency-sensitive scenarios, this delay can render state information obsolete, leading to suboptimal decisions based on outdated data and consequently degrading overall system performance [11, 19]. For instance, routing incoming user requests without real-time knowledge of server availability can cause load imbalance and increase user wait times.

Entanglement offers a fundamentally a new approach to coordination. It provides a mechanism for establishing correlations between spatially separate systems that are stronger than any achievable classically without communication [2, 3]. This mechanism can be implemented as follows. Initially, the coordinating parties share an entangled quantum state. At a later time, upon receiving local information relevant to their coordination task, each party performs a measurement on their component of the entangled state, which can be conditioned on the local information they received. The outcomes of these local measurements will exhibit strong non-local correlations and can be used to guide the parties' decisions, thus enabling coordination without communication.

In this work, we investigate the application of entanglement-assisted coordination to a dual-objective optimization problem, motivated by the scheduling challenges mentioned above. We consider a distributed system with two servers that must process both a continuously available, preemptible baseline task as well as customer requests. The goal is to maximize the throughput derived from the baseline task, subject to a Quality of Service (QoS) constraint on the average waiting time experienced by the customer requests. The challenge lies in coordinating the assignment of incoming requests to the servers based only on local information, namely the processing time required by the local request, without communication.

We develop an analytical model of this system and show that leveraging shared entanglement enables coordination strategies that outperform communication-free classical strategies. Specifically, when the baseline task's throughput increases sufficiently with processing time, entanglement allows the system to satisfy the QoS constraint while achieving higher baseline throughput. The only quantum resources required are the generation and local measurement of bipartite entanglement between the coordinating nodes. Heralded entanglement generation between physically separated systems has been demonstrated in multiple qubit platforms [6, 12, 14], including over deployed fiber [15]. This makes entanglement-assisted coordination an attractive near-term application of quantum networks.

2 RELATED WORK

The underlying principle behind leveraging entanglement for distributed coordination can be traced back to Bell's theorem. Bell demonstrated that quantum mechanics predicts correlations between spatially separated systems that are fundamentally stronger than those permitted by any classical theory based on local realism [2, 10]. This phenomenon can also be studied through the framework of non-local games, where non-communicating players cooperate to win a game against a referee. For some such games, quantum strategies allow players to achieve higher success probabilities than is classically possible without communication [3, 5].

Building on this framework, the possibility of translating the quantum advantage observed in abstract non-local games into practical benefits has been explored. The idea is to map communication or coordination problems onto the structure of a non-local game where quantum strategies are known to outperform classical ones. Notable examples of domains for which this mapping has been attempted include market making in high-frequency trading [7], load balancing in ad-hoc networks [9], and rendezvous tasks [13, 17, 18]. Our work introduces a novel application domain for entanglement-assisted coordination: optimization of scheduling in distributed systems. Further, we provide a rigorous analytical quantification of the quantum advantage in this setting.

From a different perspective, the problem studied in this work falls within the broad classical domains of load balancing, scheduling theory, and queueing theory. These fields all have extensive literature on resource allocation, task scheduling, and throughput optimization in distributed computing and networked systems across a wide range of assumptions and conditions. Foundational concepts can be found in references [4, 8, 19]. Our work differs from classical approaches by introducing entanglement as a coordination mechanism.

3 SYSTEM DESCRIPTION

We now introduce the distributed system we consider, depicted in Figure 1. It consists of two servers, represented by squares, which process work at the same rate. Each server has a queue of unlimited capacity to hold incoming requests. The system is designed to handle two distinct types of work.

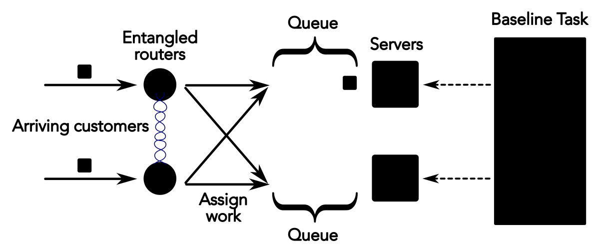

Figure 1: Distributed system studied in this work. Two servers, depicted by squares, handle two distinct types of work. On the right, a baseline task that is preemptible and always available. On the left, customer requests that arrive at the system via the routers, depicted as circles. The servers have queues of unlimited size for customers. The routers share entanglement with each other, which they use to better coordinate their routing decisions. Entanglement is depicted by wavy lines connecting the two routers.

First, there is a baseline task that is always present and available for processing by either server. We assume this baseline task is preemptible, meaning a server working on it will immediately switch to a customer request upon assignment. This assumption simplifies the model by focusing the decision-making logic entirely on the routing parties.

The second type of work consists of customer requests that arrive dynamically. We assume these requests arrive simultaneously in pairs, with one customer arriving at each of two routers (depicted as circles in Figure 1) at the same instant. This assumption was made for analytical simplicity, but we believe lifting it and timeslotting routing decisions would preserve the quantum advantage we will later observe. The customer arrival process follows a Poisson distribution with rate parameter λ. The service time required by each individual customer i is a random variable X_i drawn independently from an exponential distribution with parameter μ. This corresponds to an M/M/2 queueing system with paired arrivals. Each router observes the service time X_i required by its local customer upon arrival.

Upon the simultaneous arrival of a customer pair, the routers must assign their respective customers to one of the two servers. A key constraint of our model is that the routers cannot communicate classically in real-time to coordinate this assignment. We motivate this by considering scenarios where the physical distance between routers introduces communication latencies that are significant compared to the decision timescale and the queue dynamics, rendering decisions based on exchanged state information ineffective. Therefore, each router must make its assignment decision based solely on locally available information (i.e., its own customer's service time X_i) and any pre-shared resources or pre-established strategy.

The optimal assignment strategy depends on the combined workload. Assigning requests with a large total service time ($X_1 + X_2$) to the same server leads to long waiting times. Conversely, splitting requests with a small total service time interrupts baseline task processing on both servers. Therefore, effective coordination involves conditionally splitting or bunching requests based on the non-local information $X_1 + X_2$. This is discussed in more detail in Section 4.

The servers operate according to the following logic. If a customer is assigned to a server that is currently processing the baseline task, the task is preempted. The server then begins servicing the newly arrived customer. If a customer is assigned to a server already busy servicing a previous customer, the new customer joins that server's queue. Servers are never idle. When a server completes service for all customers in its queue and finds its queue empty, it resumes processing the baseline task until the next customer is assigned to it.

We quantify the Quality of Service (QoS) provided to customers by the average time they spend waiting in queue before service begins, denoted as W_q. We impose a strict QoS requirement by demanding that this average waiting time must not exceed a predefined limit W_l:

$$W_q \leq W_l. \tag{1}$$

The system's performance on the baseline task is quantified by its long-term average rate of output per server, denoted as $\mathbb{E}[T_B]$. We define an output function $T(t)$ that maps an uninterrupted processing time interval t during which a server works on the baseline task to the total output achieved in that interval. For a quantum advantage to manifest, $T(t)$ must exhibit sufficiently increasing returns, meaning output should grow, for example, super-linearly with processing time. This will be explained in detail in Section 4.

Combining these elements, the optimization goal for the system is to find a routing strategy, i.e., a rule for how routers assign customers based on their local information, that maximizes the long-term average baseline throughput rate $\mathbb{E}[T_B]$ while adhering to the customer waiting time constraint:

$$\text{maximize } \mathbb{E}[T_B] \quad \text{subject to } W_q \leq W_l. \tag{2}$$

The central question investigated in this work is how different types of coordination strategies, constrained by the impossibility of real-time communication, affect the achievable performance in this optimization problem. In particular, we investigate whether a coordination strategy making use of quantum resources in the form of bipartite entanglement shared by the two routers outperforms strategies that do not make use of quantum resources. In the next section we will answer this question in the affirmative.

4 MAIN RESULT

We now present the main result of this work. Entanglement can enable a higher long-term average baseline throughput rate, $\mathbb{E}[T_B]$, compared to classical strategies without communication, while respecting the same QoS constraint $W_q \leq W_l$ on average customer waiting time. This advantage arises provided the baseline output function $T(t)$, which maps an uninterrupted processing interval t to the total output achieved, satisfies $2\mathbb{E}[T(L)] < (1/\mathbb{E}[L])\mathbb{E}[LT(L)]$, with L being an exponentially distributed random variable representing the length of a server idle period (with rate parameter determined by customer arrivals). As an example, this condition is met if $T(t)$ grows super-linearly with t (e.g., $T(t) = kt^c$ with $c > 1$). A formal derivation will be given in the full version of the paper.

The quantum advantage stems from the ability of entanglement-assisted strategies to better approximate an optimal routing strategy, which is unattainable without communication. Specifically, by mapping the routers' coordinated decision problem onto a non-local game with a quantum advantage, we achieve a closer approximation to the optimal routing strategy than is classically possible.

4.1 The Optimal Threshold Strategy

The threshold strategy is defined by a threshold time t^*. Upon arrival of a customer pair with service times (X_1, X_2), if their sum $X_1+X_2 > t^*$, the pair is split; otherwise ($X_1+X_2 \leq t^*$), it is bunched. The value of t^* is chosen to meet the QoS constraint. This strategy is optimal for the optimization problem we defined.

Theorem 4.1 (Optimality of the Threshold Strategy). *For a system with Poisson customer pair arrivals (rate λ) and* *independent, exponentially distributed individual customer service times X_i (mean $1/\mu$), if the overall server utilization $\rho = \lambda/\mu < 1$, the routing strategy that maximizes the long-term average baseline throughput rate $\mathbb{E}[T_B]$ subject to $W_q \leq W_l$ is the threshold strategy.*

A detailed proof will be presented in a full version of this paper; here we provide the core intuition.

The long-term average baseline throughput rate can be expressed as $\mathbb{E}[T_B] = (1 - \rho) \cdot g(p_{\text{split}})$, where ρ is the (constant) server utilization and $g(p_{\text{split}})$ is a function of p_{split} reflecting the output generated during idle periods. If the baseline output function $T(t)$ satisfies the condition $2\mathbb{E}[T(L)] < (1/\mathbb{E}[L])\mathbb{E}[LT(L)]$ (where L is the length of an idle period, whose distribution itself depends on p_{split} via the effective arrival rate $\lambda_{eff} = \lambda(1 + p_{\text{split}})/2$), then $\mathbb{E}[T_B]$ is a strictly decreasing function of p_{split}. Under this condition on $T(t)$, maximizing $\mathbb{E}[T_B]$ is equivalent to minimizing p_{split}. The optimization problem is then finding the strategy achieving the minimum possible p_{split} while satisfying $W_q \leq W_l$.

The threshold strategy achieves this minimum p_{split}. This is because for any given p_{split} the threshold strategy with t^* appropriately chosen to yield that p_{split} minimizes the customer waiting time W_q. To see this, consider that any routing strategy will exhibit an overall long-run probability, p_{split}, that an incoming customer pair is split between the two servers (conversely, $1 - p_{\text{split}}$ is the probability it is bunched onto a single server). Any strategy with the same overall p_{split} which is not the threshold strategy must bunch some pairs with $X_1 + X_2 > t^*$, and split some with $X_1 + X_2 < t^*$. As the average waiting time is proportional to the second moment of the service time distribution, this results in a higher waiting time for the same p_{split}. Alternatively, it results in the same waiting time for a higher p_{split}, and hence lower throughput.

Perfect implementation of this strategy requires non-local knowledge of $X_1 + X_2$, making it unattainable without real-time communication. Communication-free strategies must approximate it using only local information (X_1 or X_2) and possibly pre-shared resources.

4.2 Mapping the Coordination Task to a Non-Local Game

We now go into more detail regarding how to map the coordination task described in Section 3 to a non-local game, with the routers acting as the players, A and B.

The inputs to the game are derived from the routers' observed service times X_1, X_2, which are independent random variables drawn from an exponential distribution with rate parameter μ. We transform these unbounded inputs into variables a, b uniformly distributed in the interval $[0, 1)$ using the probability integral transform via the exponential cumulative distribution function (CDF), $F(x) = 1 - e^{-\mu x}$. Specifically, we set $a = F(X_1)$ and $b = F(X_2)$. The inverse transformation is $X_i = -\frac{1}{\mu}\ln(1 - F(X_i))$.

The outputs of the game, $o_A, o_B \in \{+1, -1\}$, correspond to the routing decision. We map the decision to bunch customers to the output condition $o_A \cdot o_B = +1$, and the decision to split customers to $o_A \cdot o_B = -1$.

The winning condition of the game reflects the optimal threshold strategy for the original routing problem. Players win if their output decision matches the ideal decision based on the total service time $X_1 + X_2$ relative to a threshold t. The condition to bunch, $X_1 + X_2 \leq t$, transforms under the CDF mapping as follows:

$$-\frac{1}{\mu}\ln(1-a) - \frac{1}{\mu}\ln(1-b) \leq t \implies (1-a)(1-b) \geq e^{-\mu t}. \quad (3)$$

Similarly, the condition to split, $X_1 + X_2 > t$, transforms to $(1 - a)(1 - b) < e^{-\mu t}$.

Therefore, the coordination task is equivalent to a non-local game where players receive inputs $a, b \sim U(0, 1)$ and win if their outputs o_A, o_B satisfy:

$$o_A \cdot o_B = \begin{cases} +1 & \text{if } (1-a)(1-b) \geq C_t \\ -1 & \text{if } (1-a)(1-b) < C_t \end{cases}, \quad (4)$$

where $C_t = e^{-\mu t}$ is a constant determined by the threshold t and rate parameter μ. The objective is to find strategies $o_A(a)$ and $o_B(b)$ that maximize the probability P_{win} of satisfying this condition, averaged over the uniform input distributions. Note that P_{win} is both the probability of winning the non-local game and the probability of correctly implementing the ideal threshold strategy.

This game shares structural similarities with previously studied continuous-input non-local games, in particular the first of three games introduced in [1], namely the use of independent uniform inputs $a, b \in [0, 1)$ and binary outputs $o_A, o_B \in \{+1, -1\}$ aiming for correlation without communication. However, the boundary defining the winning condition is different. While the game in [1] features a linear boundary ($a + b = 1$), our game derived from the routing problem features the hyperbolic boundary $(1 - a)(1 - b) = C_t$. Due to this difference, the known optimal classical ($P_{\text{win}} = 0.75$) and quantum ($P_{\text{win}} \approx 0.818$) winning probabilities for the game in [1] do not directly apply to our scenario. Nonetheless, the proven existence of a quantum advantage in that game motivates exploring quantum strategies for the game derived here. In the subsequent sections, we will investigate the maximum achievable P_{win} for both classical and entanglement-assisted quantum strategies for the game we have defined using numerical optimization techniques.

4.3 Classical Strategy

Determining the optimal classical strategy that maximizes the winning probability $P_{\text{win}}^{\text{classical}}$ for the non-local game defined in (4) is analytically challenging, with the difficulty arising from the hyperbolic boundary $(1-a)(1-b) = C_t$ defining the winning condition.

Necessary conditions for optimality in non-local games often lead to systems of differential equations [16]. For example, a non-local game with a linear boundary [1] yields a simple system of differential equations whose solutions can be fully characterized, thus allowing for rigorous proofs of the optimal classical and quantum strategies [16]. However, applying the same methods to our non-local game results in a much more complex system of coupled differential-functional equations that we have not been able to solve analytically. This prevents a proof of the optimal classical strategy and its corresponding maximum winning probability.

We hence turn to numerical optimization. We start by noting that deterministic strategies are known to be optimal for non-local games among the set of non-communicating classical strategies [3], and we can thus restrict our optimization to such strategies. We discretize the continuous input intervals $[0, 1)$ for a and b into N bins (we used $N = 1000$). Classical deterministic strategies are then represented by vectors $\mathbf{G}, \mathbf{H} \in \{+1, -1\}^N$, corresponding to piecewise constant functions $g(a)$ and $h(b)$. The objective is to find the vectors $\mathbf{G}, \mathbf{H}$ that maximize the expected agreement between the players' output product $o_A o_B = g(a)h(b)$ and the ideal outcome specified by the game rule, averaged over the uniform input distribution. The maximum winning probability is then given by $P_{\text{win}}^{\text{classical}} = (1 + \max_{\mathbf{G},\mathbf{H}} \mathbb{E}[o_A o_B \cdot \text{Rule}])/2$.

We employ simulated annealing to search the large (2^{2N}) space of possible deterministic strategy vectors $(\mathbf{G}, \mathbf{H})$. This optimization is performed for various values of the threshold constant C_t, so as to find the value of C_t for which the maximum $P_{\text{win}}^{\text{classical}}$ is minimized. We expect that this represents the 'hardest' case for classical strategies and thus offers the largest potential margin for quantum advantage. We find that the minimum occurs at $C_t \approx 0.33$, yielding:

$$P_{\text{win}}^{\text{classical}}(C_t \approx 0.33) \approx 0.7616. \quad (5)$$

We use this value as the classical benchmark, although we acknowledge that we have no formal proof of its optimality.

4.4 Quantum Strategy

To potentially outperform the classical benchmark found in Section 4.3, the routers can leverage pre-shared quantum entanglement. We assume the two routers share the state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Upon observing their local service times X_1, X_2, each router performs a local projective measurement on their qubit. Crucially, the choice of measurement basis for router i depends on their local input X_i. This is implemented by mapping the input X_i to a measurement angle θ_i, using the transformed variable $F(X_i) = 1 - e^{-\mu X_i}$ obtained via the probability integral transform, discussed in Section 4.2. This angle θ_i defines the local measurement basis in the XY-plane of the Bloch sphere, relative to the computational basis. One pair of entangled qubits is consumed per routing decision.

We employ numerical optimization to find the angle functions $\theta_A(F(X_1))$ and $\theta_B(F(X_2))$ maximizing the quantum winning probability $P_{\text{win}}^{\text{quantum}}$. The winning probability is calculated by integrating probability of obtaining the desired outcome (same or different, according to the game rule) multiplied by the joint probability density function $\mu^2 e^{-\mu(X_1+X_2)}$ over $X_1, X_2 \in [0, \infty)$.

We explored two parameterizations: We parametrized $\Delta = \theta_A(F(X_1)) - \theta_B(F(X_2))$ as depending linearly on the inputs $F(X_1), F(X_2)$, i.e., $\Delta = d_0 + d_1 F(X_1) + d_2 F(X_2)$. The optimization yielded $d_0 \approx 3.709$, $d_1 \approx -2.760$, and $d_2 \approx -2.759$, resulting in a maximum quantum winning probability of:

$$P_{\text{win, linear}}^{\text{quantum}}(C_t \approx 0.33) \approx 0.8140. \quad (6)$$

We tried also allowing for a quadratic term, which slightly improved this to $P_{\text{win, quad}}^{\text{quantum}}(C_t \approx 0.33) \approx 0.8147$. This minor increase suggests that the optimal strategy is well-approximated by the linear relative angle model.

4.5 Quantum Advantage

We now argue why $P_{\text{win}}^{\text{quantum}} > P_{\text{win}}^{\text{classical}}$ translates to improved performance on the optimization problem. As per Section 4.1 any deviation from the ideal threshold strategy leads to an increase in the average waiting time W_q. Therefore, to satisfy a given QoS constraint $W_q \leq W_l$, a strategy that deviates more from the ideal threshold decisions must operate at a higher overall splitting probability p_{split} to compensate.

Let p_{ideal} be the splitting probability corresponding to the threshold strategy that exactly meets the constraint $W_q = W_l$. A non-communicating classical strategy, being less accurate in approximating the threshold strategy, will effectively deviate more often. Let p_c and p_q be the splitting probabilities required for the classical and quantum strategies to satisfy the constraint $W_q \leq W_l$. As both quantum and classical deviate from the ideal strategy, with quantum deviating less, it holds that $p_{ideal} < p_q < p_c$.

If the baseline output function $T(t)$ satisfies the condition $2\mathbb{E}[T(L)] < (1/\mathbb{E}[L])\mathbb{E}[LT(L)]$, the long-term average baseline throughput rate $\mathbb{E}[T_B]$ is a strictly decreasing function of p_{split}. Combining these arguments leads to our main result: because the quantum strategy achieves $p_q < p_c$ while meeting the same QoS, it yields a higher baseline throughput rate:

$$\mathbb{E}[T_B]_{\text{quantum}} > \mathbb{E}[T_B]_{\text{classical}}. \quad (7)$$

4.6 Example: Polynomial Throughput Functions

As an example, we now quantify the quantum advantage for baseline tasks where the total output function $T(t)$ is a monomial: $T(t) = ct^k$ for $t \geq 0$, where $c > 0$ is a scaling constant and $k \geq 0$ is the degree. We set $c = 1$ without loss of generality for analyzing relative performance.

The long-term average baseline throughput rate is $\mathbb{E}[T_B] = (1 - \rho_{\text{const}}) \cdot g(p_{\text{split}})$, where $\rho_{\text{const}} = \lambda E[X_i]$ is constant and $g(p_{\text{split}}) = \lambda_{\text{eff}}(p_{\text{split}})\mathbb{E}_{L \sim Exp(\lambda_{\text{eff}}(p_{\text{split}}))}[T(L)]$. For $T(t) = t^k$, this simplifies to $\mathbb{E}[T_B] = (1 - \rho_{\text{const}})\frac{k!}{(\lambda_{\text{eff}}(p_{\text{split}}))^{k-1}}$, where $\lambda_{\text{eff}}(p_{\text{split}}) = \lambda(1 + p_{\text{split}})/2$.

For $\mathbb{E}[T_B]$ to be strictly decreasing with p_{split} the exponent $k - 1$ must be positive, implying $k > 1$. This condition signifies strictly increasing returns: longer uninterrupted processing periods become disproportionately more valuable. If $k = 1$ (linear returns, $T(t) = t$), $\mathbb{E}[T_B]$ is independent of p_{split}, yielding no throughput advantage from a lower p_{split}. We focus on the $k > 1$ case.

The ratio R_k of baseline throughput rates achieved by a quantum strategy (with splitting probability p_q) and a classical strategy (with p_c), where $p_q < p_c$ for the same QoS, is given by:

$$R_k = \left(\frac{1 + p_c}{1 + p_q}\right)^{k-1}. \quad (8)$$

The derivation will be given in the long version of this paper. Since $p_c > p_q$ for the same QoS requirement, and assuming $k > 1$, we have $R_k > 1$. The magnitude of this advantage, R_k, grows with as a power of the degree $k - 1$. This demonstrates that even modest gains in raw coordination accuracy can translate into substantial performance benefits for baseline tasks exhibiting strong increasing returns.

5 CONCLUSION

We have shown that shared entanglement can significantly improve coordination in distributed scheduling problems, leading to higher system throughput compared to classical strategies without real-time communication. This advantage is most pronounced for tasks with increasing returns, where even modest gains in coordination accuracy, achieved by mapping the routing problem onto a non-local game with a quantum advantage, translate into substantial performance benefits. Our results highlight distributed scheduling as a promising, practical application domain for near-term quantum networks, given that the required resources amount to the generation and local measurement of bipartite entanglement, a capability that has been experimentally demonstrated with multiple physical systems [6, 12, 14], including over deployed fiber [15].

The assumption of simultaneous arrival of customer pairs is somewhat unnatural. While this assumption made analysis simpler, we believe it could be lifted by timeslotting routing decisions without losing the underlying quantum advantage. Furthermore, while our numerical evidence for quantum advantage is strong for the non-local game derived from exponential service times we introduced (for which rigorous bounds are still open), the core benefit stems from applying non-local quantum correlations to a coordination task constrained by locality. This fundamental principle suggests that a quantum advantage should persist for other service time distributions, even though they would map to different non-local games with potentially different optimal strategies and advantage magnitudes.

While the simple system we studied in this work is of independent interest, an obvious next step is to identify specific scenarios within practical domains that map effectively onto the core principles of the model studied here. These would then constitute practical use cases for entanglement-based quantum networks. Promising candidates for further exploration include content-delivery networks and medium-access control in wireless networking.

The identification of this novel application domain gives rise to multiple promising avenues for further research. These include investigating connections to, and potential enhancements for, established load balancing and scheduling algorithms from the classical literature by incorporating entanglement resources; exploring the applicability of other non-local games or quantum correlation protocols to different coordination problems; and extending the analysis to scenarios involving more than two coordinating nodes, potentially leveraging multipartite entanglement or networks of bipartite links.

This work does not raise any ethical issues.

REFERENCES

[1] Netanel Aharon, Shai Machnes, Benni Reznik, Jonathan Silman, and Lev Vaidman. 2013. Continuous input nonlocal games. *Natural computing* 12 (2013), 5–8.

[2] John S Bell. 1964. On the einstein podolsky rosen paradox. *Physics Physique Fizika* 1, 3 (1964), 195.

[3] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. 2014. Bell nonlocality. *Reviews of modern physics* 86, 2 (2014), 419–478.

[4] Thomas L. Casavant and Jon G. Kuhl. 1988. A taxonomy of scheduling in general-purpose distributed computing systems. *IEEE Transactions on software engineering* 14, 2 (1988), 141–154.

[5] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. 1969. Proposed experiment to test local hidden-variable theories. *Physical review letters* 23, 15 (1969), 880.

[6] Jacob P Covey, Harald Weinfurter, and Hannes Bernien. 2023. Quantum networks with neutral atom processing nodes. *npj Quantum Information* 9, 1 (2023), 90.

[7] Dawei Ding and Liang Jiang. 2024. Coordinating Decisions via Quantum Telepathy. *arXiv preprint arXiv:2407.21723* (2024).

[8] Donald Gross, John F Shortle, James M Thompson, and Carl M Harris. 2011. *Fundamentals of queueing theory*. Vol. 627. John wiley & sons.

[9] Masoud Hasanpour, Shahin Shariat, Payam Barnaghi, Seyed Amir Hoseinitabatabaei, Seiamak Vahid, and Rahim Tafazolli. 2017. Quantum load balancing in ad hoc networks. *Quantum Information Processing* 16 (2017), 1–29.

[10] Bas Hensen, Hannes Bernien, Anaïs E Dréau, Andreas Reiserer, Norbert Kalb, Machiel S Blok, Just Ruitenberg, Raymond FL Vermeulen, Raymond N Schouten, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Matthew Markham, Daniel J. Twitchen, David Elkouss, Stephanie Wehner, Tim H. Taminiau, and Ronald Hanson. 2015. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* 526, 7575 (2015), 682–686.

[11] Yichuan Jiang. 2015. A survey of task allocation and load balancing in distributed systems. *IEEE Transactions on Parallel and Distributed Systems* 27, 2 (2015), 585–599.

[12] V. Krutyanskiy, M. Galli, V. Krcmarsky, S. Baier, D. A. Fioretto, Y. Pu, A. Mazloom, P. Sekatski, M. Canteri, M. Teller, J. Schupp, J. Bate, M. Meraner, N. Sangouard, B. P. Lanyon, and T. E. Northup. 2023. Entanglement of Trapped-Ion Qubits Separated by 230 Meters. *Phys. Rev. Lett.* 130 (Feb 2023), 050803. Issue 5. https://doi.org/10.1103/PhysRevLett.130.050803

[13] Piotr Mironowicz. 2023. Entangled rendezvous: a possible application of Bell non-locality for mobile agents on networks. *New Journal of Physics* 25, 1 (2023), 013023.

[14] Maximilian Ruf, Noel H Wan, Hyeongrak Choi, Dirk Englund, and Ronald Hanson. 2021. Quantum networks based on color centers in diamond. *Journal of Applied Physics* 130, 7 (2021).

[15] Arian J. Stolk, Kian L. van der Enden, Marie-Christine Slater, Ingmar te Raa-Derckx, Pieter Botma, Joris van Rantwijk, J. J. Benjamin Biemond, Ronald A. J. Hagen, Rodolf W. Herfst, Wouter D. Koek, Adrianus J. H. Meskers, René Vollmer, Erwin J. van Zwet, Matthew Markham, Andrew M. Edmonds, J. Fabian Geus, Florian Elsen, Bernd Jungbluth, Constantin Haefner, Christoph Tresp, Jürgen Stuhler, Stephan Ritter, and Ronald Hanson. 2024. Metropolitan-scale heralded entanglement of solid-state qubits. *Science Advances* 10, 44 (2024), eadp6442. https://doi.org/10.1126/sciadv.adp6442 arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.adp6442

[16] Boris Tsirelson. 2007. Some extremal problems related to Bell-type inequalities. *arXiv preprint arXiv:0706.1091* (2007).

[17] J Tucker, Paul Strange, Piotr Mironowicz, and Jorge Quintanilla. 2024. Quantum-assisted Rendezvous on Graphs: Explicit Algorithms and Quantum Computer Simulations. *arXiv preprint arXiv:2405.14951* (2024).

[18] Giuseppe Viola and Piotr Mironowicz. 2024. Quantum strategies for rendezvous and domination tasks on graphs with mobile agents. *Physical Review A* 109, 4 (2024), 042201.

[19] Yung-Terng Wang and Robert J. T. Morris. 1985. Load sharing in distributed systems. *IEEE Transactions on computers* 100, 3 (1985), 204–217.