Exploring the Structural Property of the Optimal Entanglement Policy for Quantum Switch

Bin Luo*, Xiaojun Lin, John C.S. Lui
The Chinese University of Hong Kong
{bluo23, cslui}@cse.cuhk.edu.hk, xjlin@ie.cuhk.edu.hk

***Abstract*—A quantum switch is one of the most fundamental network elements for connecting different quantum devices. In this paper, we explore the "*optimal entanglement policy*" of a quantum switch under a scenario where the stored and entangled qubits in the quantum switch may undergo a decoherence process. Finding an optimal entanglement policy is important as it enables a quantum switch to make a judicious basis measurement based on the number of existing link-level entanglements to maximize the "*weighted throughput*". We use a Markov decision process framework to model the dynamics of the quantum switch, and theoretically shows a *threshold-based property* between bipartite and tripartite policies under a very general class of weight functions with a weight parameter β. Empirically, such a *threshold-based property* also holds for the optimal entanglement policy. In particular, the quantum switch should perform a bipartite or tripartite policy when β is below a threshold β_B^* or above a threshold β_T^*, respectively. When $\beta_B^* < \beta < \beta_T^*$, the quantum switch needs to perform a "*threshold-based and state-dependent entanglement policy*". We also extend the work to allow a mixture of bipartite and tripartite policies. We theoretically show similar threshold-related relationships between mixed and bipartite/tripartite policies. We carry out extensive numerical experiments to confirm that the optimal entanglement policy has such structural property.**

I. Introduction and Related Work

Quantum computing uses quantum physics principles, such as superposition and entanglement [1], enabling quantum computers to achieve far greater performance than classical computers, with potentially transformative impacts in fields such as communication [2]–[6], cryptography [7] and chemical simulations [8]. A quantum network is a distributed infrastructure that connects different quantum devices to facilitate the exchange of information [9]–[13]. Similar to the classical Internet, a quantum network can harness the power of distributed quantum computation [14]–[16]. Among the different quantum devices in a quantum network, one important device is the quantum switch. The quantum switch first establishes *link-level entanglements* between itself and its connected users. Then, as these link-level entanglements emerge, the quantum switch can perform bipartite basis measurements between pairs of locally-held qubits or tripartite basis measurements among triples of locally-held qubits, thereby generating *bipartite or tripartite end-to-end user entanglement*. These entanglements are instrumental for various quantum computing operations, including quantum key distribution [17]–[19] and beyond.

In this work, we are interested in policies that can maximize the throughput for providing such bipartite and tripartite end-to-end user entanglements. We define the bipartite/tripartite entanglement throughput as the rate at which a quantum switch provides bipartite/tripartite user entanglement under a given policy. The "*total throughput*" of a quantum switch is simply the sum of bipartite and tripartite entanglement throughput. We further extend this concept to the "*weighted (total) throughput*", which is the weighted sum of bipartite and tripartite entanglement throughput. One can easily adjust these weights to reflect the relative importance of bipartite and tripartite user entanglements for various applications [17], [18], [20]. Our goal is then to find online policies that can maximize the weighted throughput.

Due to the importance of the quantum switch in quantum networks, researchers have made efforts to study its performance. Authors in [21]–[23] used continuous-time Markov chain (CTMC) and discrete-time Markov chain (DTMC) to model the quantum switch and provide closed-form expressions for the maximal generation rate of n-qubit end-to-end user entanglement, which is referred to as the capacity C_n, under various scenarios. However, these works assumed that the quantum switch can only perform one type of basis measurement. Hence, they did not address the question of how to optimally decide between multiple types of user entanglements (e.g., bipartite vs. tripartite). In contrast, [24] considered both bipartite ($n = 2$) and tripartite ($n = 3$) entanglements for users. It provided elegant expressions for both bipartite capacity C_2 and tripartite capacity C_3, and showed that the complete capacity region of the quantum switch is approximately equal to the capacity region under a set of time-division multiplexing (TDM) policies, as the number of connected users increases. However, these TDM policies are mainly for characterizing the capacity region, and do not provide an online decision policy. In other words, they do not indicate how to control the bipartite/tripartite entanglements to maximize the weighted throughput. In the literature, online decision algorithms have been proposed in [25]–[30] to determine an entanglement policy which can stabilize the quantum switch. However, these studies either did not consider decoherence [25], [29], or imposed the restrictive assumption that qubits survive only one unit of time in a discrete-time setting [26]–[28], [30]. As a result, they present an either overly optimistic or overly

*Bin Luo is the corresponding author. The work of John C.S. Lui was supported in part by the RGC SRFS2122-4S02. This work of Xiaojun Lin was supported by the Government of the Hong Kong SAR through the Global STEM Professorship. We sincerely thank the reviewers for dedicating their valuable time to evaluate our work and for providing insightful suggestions.

pessimistic model that does not accurately reflect the real-world behavior of quantum systems, leading to overestimation or underestimation of performance, e.g., weighted throughput.

Thus, one open question is, *what is the optimal entanglement policy which can maximize the weighted throughput, when link-level entanglement decoherence exists in a continuous-time setting*? Specifically, we aim to answer the following question in this work: Given the number of existing link-level entanglements, which state-dependent entanglement action should the quantum switch take to maximize the weighted throughput? We answer this question by revealing an important structure of the optimal entanglement policy under the presence or absence of qubits decoherence. The contributions of this work are as follows:

- We use a Markov decision process (MDP) framework to model the dynamics of a quantum switch and theoretically prove the equivalence between the expected average reward and the weighted throughput. Using the relative value iteration algorithm, we empirically uncover the "*threshold-based structural property*" of the optimal entanglement policy, whether decoherence exists or not.
- We theoretically show the "*threshold-based property*" between two special optimal entanglement policies under a very *general class of weight functions* with a weight parameter β. Specifically, we compare a bipartite policy and tripartite policy under two settings, one with qubit decoherence and one without. In both settings, we theoretically show that the tripartite policy achieves higher weighted throughput than the bipartite policy when β exceeds a threshold β^{BT}, whereas the bipartite policy is more advantageous when β is below the threshold β^{BT}. Extensive numerical experiments are performed to substantiate our theoretical results.
- We combine the bipartite policy and the tripartite policy to create a family of "*mixed entanglement policies*". We theoretically show that, for any specific mixed policy, the threshold-based property also exists between the mixed policy and the bipartite/tripartite policy. Numerical experiment results validate our theoretical results on the threshold-based phenomenon and the dynamic interplay between the mixed policy and the bipartite/tripartite policy, with or without the presence of qubit decoherence.

II. System Model and Formulation

Model Description: As illustrated in Figure 1(a), we consider a quantum switch connected to $k \geq 3$ users. Each user possesses a quantum device connected to the quantum switch via a quantum channel/link. The quantum switch's primary role is to enable *end-to-end entanglements* among k users through a two-phase process: Initially, *link-level entanglements* are formed between the quantum switch and an individual user via a quantum channel, as depicted in red in Figure 1(b). Each link-level entanglement forms a Bell state, which utilizes two qubits: one within the quantum switch and the other within the user's quantum device. Subsequently, upon generating $2 \leq n \leq k$ link-level entanglements across distinct quantum channels (Figure 1(c)), the switch executes a quantum swapping operation to establish n-qubit end-to-end user entanglement (Figure 1(d)). This involves basis measurements on the stored qubits within the switch: either a two-qubit Bell-state measurement for $n = 2$ (Figure 1(e)), or an n-qubit GHZ-state measurement for $n \geq 3$, akin to Figure 1(f). Given the prevalence of two-qubit and three-qubit entangled states in quantum communication systems [31]–[33], we assume that the quantum switch can perform either bipartite ($n = 2$) or tripartite ($n = 3$) basis measurement.

Unlike a classical switch where events are deterministic [34], the success of establishing the link-level entanglements and subsequent basis measurements in a quantum switch are probabilistic. Assuming homogeneous[1] and statistically independent links, the arrivals of attempts to establish link-level entanglements on each channel are modeled as a Poisson process[2] with a mean rate τ, and each arrival achieves success with probability p. Thus, the arrival process of successful establishment of link-level entanglements on a channel is represented as a Poisson process with a mean rate $\lambda = p\tau$. Furthermore, we assume that the execution rate of bipartite basis measurements by the quantum switch is a Poisson process characterized by a mean rate τ_2, and with a successful measurement probability of p_2. As a result, this procedure's successful completion is modeled as a Poisson process with a mean rate $\mu_2 = p_2\tau_2$. The procedure for successful tripartite basis measurements follows a similar pattern, with a mean rate $\mu_3 = p_3\tau_3$. Since performing a tripartite basis measurement is more difficult than bipartite basis measurement [37], i.e., the successful probabilities and the attempting rates satisfies $p_3 < p_2$ and $\tau_3 < \tau_2$, it is inherently easier for the quantum switch to provide bipartite user-entanglement than tripartite user-entanglement, i.e., $\mu_3 < \mu_2$.

After a link-level entanglement is successfully created, but before it is consumed by either bipartite or tripartite basis measurement, the corresponding qubits are preserved in quantum memories. In this work, we assume that each link only has one buffer[3], i.e., $B = 1$, meaning a maximum of one link-level entanglement can be stored per link. Due to environmental susceptibility of stored qubits, these entanglements will undergo decoherence, resulting in low fidelity and rendering them useless for user connections. We model the time interval to discard a link-level entanglement with low fidelity as a random variable following an exponential distribution with mean $1/\alpha$. Since the value of α depends on the lifetime of stored qubits, we refer to α as decoherence-associated memory cut-off rate [24] or simply the decoherence rate, where $\alpha = 0$ implies the scenario of no decoherence. In this work, we do not study how decoherence and basis measurements affect the fidelity of resulting user entanglements. Instead, we focus

[1]Links are homogeneous/heterogeneous if the arrival rates of link-level entanglement on each link are the same/different.

[2]Using Poisson processes to model the dynamics of a quantum switch offer benefits in incorporating the stochastic occurrences of qubit decoherence [23], [24], [35] and aligns with the Markov property required in MDPs [36].

[3]Assuming $B = 1$ reflects the current technological constraints in quantum memory [38]. For $B > 1$, it is in our future work.

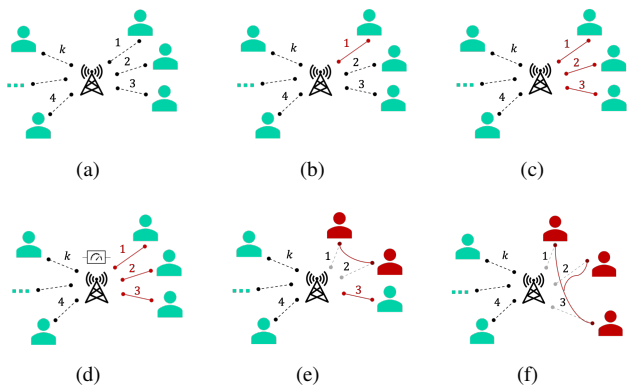


Fig. 1: Dynamics of a quantum switch. (a) The quantum switch connects k users via separate and homogeneous links. Initially, no link-level entanglement exists. (b) A link-level entanglement is generated on one of the k links (solid line), with one qubit stored in the quantum switch and the other at the user's device. (c) Over time, three link-level entanglements are generated on three different links. (d) The quantum switch can perform either bipartite or tripartite basis measurements. (e) If a bipartite basis measurement is successfully performed, the quantum switch provides a bipartite end-to-end user entanglement for 2 users (in red icon). (f) If a tripartite basis measurement is successfully performed, the quantum switch provides a tripartite end-to-end user entanglement for 3 users (in red icon).

on their effects on the structure of the optimal entanglement policy and the weighted throughput of the quantum switch.

MDP Formulation: As the network accrues link-level entanglements across various quantum channels, the quantum switch faces a strategic decision: (1) *wait for a new link-level entanglement arrival*, or (2) *perform bipartite basis measurement*, or (3) *perform tripartite basis measurement*. For instance, with two existing link-level entanglements, the quantum switch can choose to wait for a new link-level entanglement arrival to perform tripartite basis measurement which connects three users. However, this incurs delay due to the waiting of a new link-level entanglement and may inherently lower success probability of tripartite basis measurement compared to bipartite basis measurements. Moreover, such waiting may increase the chance that the entangled links become useless in the presence of qubit decoherence.

To find the right decision between bipartite and tripartite basis measurements in maximizing the weighted throughput, we use the theory of MDP [39]–[42] to model the dynamics of the quantum switch and explore the optimal entanglement policy to establish end-to-end user entanglements enabling the quantum switch to achieve maximal weighted throughput. Our MDP formulation of the quantum switch is as follows:

• **State:** The state of the quantum switch is defined by the number of existing link-level entanglements. Formally, its state space is $\mathcal{S} = \{S_0, S_1, \cdots, S_k\}$ for k connected users, where state S_i indicates that there are $0 \leq i \leq k$ existing link-level entanglements.

• **Action:** The action space is defined as $\mathcal{A} = \{A_0, A_2, A_3\}$, where A_0 corresponds to the quantum switch's decision of not doing any basis measurement but wait for further link-level entanglement arrivals. A_2 corresponds to performing the bipartite basis measurements. Besides, A_3 corresponds to performing the tripartite basis measurements.

• **Reward:** For action A_0, the reward $r(S_i, A_0), 0 \leq i \leq k$, is set to zero, as the quantum switch's inaction yields no generation of user entanglement. For action A_2, we define the reward $r(S_i, A_2) = w_B \mu_2 1\{i \geq 2\}$, where $w_B > 0$ is the weight assigned to the bipartite entanglement throughput, indicating the importance of the bipartite user entanglement to the overall system performance. The reward of A_3 is defined as $r(S_i, A_3) = w_T \mu_3 1\{i \geq 3\}$, where $w_T > 0$ is the weight assigned to the tripartite entanglement throughput.

• **Transition probability matrix:** Following the dynamics of a quantum switch, one can construct a $|\mathcal{S}|$-by-$|\mathcal{S}|$ transition rate matrix Q^a for each action $a \in \mathcal{A}$. Given Q^a as a rate matrix under consistent action a across all states $s \in \mathcal{S}$, one can use the uniformization technique [43] [44] to transform the CTMC to a DTMC with a transition probability matrix $P^a = I + \frac{1}{\rho} Q^a$. Here, ρ is the uniformization rate parameter exceeding or equal to the maximum absolute diagonal value of Q^a for every action $a \in \mathcal{A}$. Consequently, the MDP's transition probability matrix $\mathbb{P}$ is derived as $\mathbb{P}(S_j | S_i, a) = P^a(i, j)$, representing the probability that the switch change from state S_i to S_j under action a. See the Appendix A for the detailed construction of the transition probability matrix $\mathbb{P}$.

To simplify notations, we define the expectation operator $P_t^\pi, \forall t \in \mathbb{N}$, which satisfies $P_0^\pi = I$ and $P_{t+1}^\pi = P^\pi P_t^\pi$, where $P^\pi(i, j) = \mathbb{P}(S_j | S_i, \pi(S_i))$, $0 \leq i, j \leq k$. When the operator P^π applies to a function $f(i)$, it satisfies $P^\pi f(i) = \sum_{j=0}^{k} P^\pi(i, j) f(j), 0 \leq i \leq k$. Besides, we denote $r(S_i, \pi(S_i))$ as $r^\pi(i), 0 \leq i \leq k$, for simplicity. For a given policy $\pi : \mathcal{S} \to \mathcal{A}$, let $V_N^\pi(i)$ denotes the total expected reward over $N \geq 0$ periods under the policy π, starting from state S_i. Initially, $V_0^\pi(i) = 0$ for every $0 \leq i \leq k$. Formally, $\forall 0 \leq i \leq k, N \geq 1$, it holds that

$$V_N^\pi(i) = r^\pi(i) + P^\pi V_{N-1}^\pi(i) = \sum_{t=0}^{N-1} P_t^\pi r^\pi(i). \quad (1)$$

Then $g^\pi(i) = \lim_{N \to \infty} \frac{1}{N} V_N^\pi(i)$ is the expected average reward of the policy π starting from state S_i. If the policy π induces an irreducible Markov chain with finite states, then the expected average reward g^π is the same for all initial states satisfying

$$g^\pi = \lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} \sum_{j=0}^{k} P_t^\pi(i, j) r^\pi(j) = \sum_{j=0}^{k} \Pi(j) r^\pi(j), \quad (2)$$

where Π is the stationary distribution on $\mathcal{S}$, and $\Pi(j)$ represents the average fraction of time staying in state S_j [36].

• **Optimization goal:** Based on the definition of the reward, the expected average reward g^π can be written as

$$g^\pi = w_B \sum_{\pi(S_j) = A_2} \Pi(j) \mu_2 + w_T \sum_{\pi(S_j) = A_3} \Pi(j) \mu_3, \quad (3)$$

$$= w_B C_B^\pi + w_T C_T^\pi, \quad (4)$$

where C_B^π and C_T^π are bipartite and tripartite entanglement throughput under the policy π, respectively. Note that when $w_B = 1$ and $w_T = 1$, g^π represents the *total throughput* of the quantum switch. Our optimization goal is to determine the optimal policy π that maximizes the weighted throughput g^π. Mathematically, it can be written as

$$\max_\pi g^\pi = \lim_{N\to\infty} \frac{1}{N} V_N^\pi(i). \tag{5}$$

For this optimization goal, we can obtain the optimal policy by using the relative value iteration algorithm [36].

III. Theoretical Analysis

A. Bipartite Policy vs. Tripartite policy

As described in Section II, the quantum switch needs to make a strategic decision in selecting the appropriate basis measurement to maximize the weighted throughput, given the current number of link-level entanglements. In this section, we first prove *the existence of a threshold-based criterion* which governs the choice between bipartite and tripartite policies, with or without presence of qubit decoherence.

Specifically, we are interested in conditions when using purely bipartite or tripartite basis measurements is a better choice. If a policy always uses bipartite basis measurement whenever there is a minimum of two link-level entanglements. we refer to it as the bipartite policy π^B, defined mathematically as $\pi^B(S_i) = A_0 1\{0 \le i \le 1\} + A_2 1\{2 \le i \le k\}$. The reward for each state S_i under the bipartite policy π^B is defined as $r^B(i) = r(S_i, \pi^B(S_i)) = w_B \mu_2 1\{2 \le i \le k\}$. The transition dynamics under the bipartite policy are illustrated in Figure 2. Let P^B denotes the transition probability matrix of the Markov chain under policy π^B. For brevity, the total expected reward $V_N^{\pi^B}(i)$ under policy π^B, as per Eq.1, is denoted as $V_N^B(i)$, $0 \le i \le k$, while the expected average reward g^{π^B} defined in Eq.2 is denoted as g^B. Since bipartite policy π^B induces an irreducible Markov chain with finite states, g^B must exist [36].

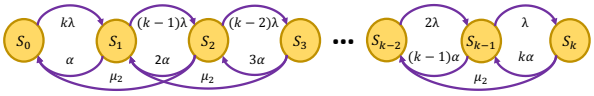


Fig. 2: The transition rate diagram of the Markov chain determined by the bipartite policy π^B.

Moreover, we consider another policy, termed as the tripartite policy π^T, wherein the quantum switch always uses tripartite basis measurement whenever there are no fewer than three link-level entanglements. Formally, this policy can be expressed as $\pi^T(S_i) = A_0 1\{0 \le i \le 2\} + A_3 1\{3 \le i \le k\}$. Accordingly, the reward under policy π^T is $r^T(i) = r(S_i, \pi^T(S_i)) = w_T \mu_3 1\{3 \le i \le k\}$. The Markov chain governed by π^T is defined by a transition probability matrix $P^T = P^{A_3}$, obtained from the original uniformization technique. The total expected reward $V_N^T(i) = V_N^{\pi^T}(i)$ and the expected average reward g^T of the policy π^T are defined in accordance to Eq.1 and Eq.2 for all $0 \le i \le k$, respectively.

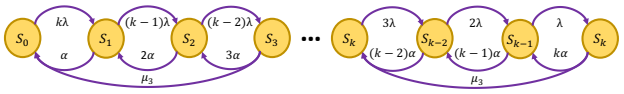


Fig. 3: The transition rate diagram of the Markov chain determined by the tripartite policy π^T.

Since the Markov chain with finite states induced by the tripartite policy π^T is irreducible, g^T must exist.

In order to analyze how varying the emphasis on bipartite and tripartite user entanglement affects the weighted throughput, we consider the weights to be *a family of weight functions* with a weight parameter $\beta \in (-\infty, +\infty)$, defined as $w_B = \eta^\beta$ and $w_T = (\eta + \Delta)^\beta$ where $\eta \in (1, +\infty)$ and $\Delta \in (0, +\infty)$. Note that this form of weight functions offers convenience in obtaining optimal policies under different scenarios. By simply adjusting the weight parameter such that $\beta > 0$ or $\beta < 0$, we can obtain the optimal policies for cases where $w_T > w_B$ or $w_T < w_B$, respectively. When $\beta = 0$, optimal policies of the optimization goal (5) maximize the total throughput.

Theorem 1. *Suppose the weight functions satisfy $w_B = \eta^\beta$, $w_T = (\eta + \Delta)^\beta$, where $\eta > 1$ and $\Delta > 0$ are constants, $\rho > \max\{\max_{0\le i\le k} |Q^B(i,i)|, \max_{0\le i\le k} |Q^T(i,i)|\}$ and deocherence rate $\alpha \ge 0$. If there exists constants $c_B, c_T \in (0,1]$ which satisfies $\sum_{j=2}^k P_t^B(i,j) \ge c_B$ and $\sum_{j=3}^k P_t^T(i,j) \ge c_T$ for every $t \ge 0$ and $0 \le i \le k$, there must exist a threshold*

$$\beta^{BT} \in [-\log_{\frac{\eta}{\eta+\Delta}}(\frac{c_B\mu_3}{\mu_2}), \log_{\frac{\eta}{\eta+\Delta}}(\frac{c_T\mu_3}{\mu_2})]$$

such that $g^B > g^T$ when $\beta < \beta^{BT}$, and $g^B < g^T$ when $\beta > \beta^{BT}$.

Proof sketch.[4] For any $N \ge 0$, we can deduce that $V_N^B(i) - V_N^T(i) = (\eta + \Delta)^\beta f_i^N(\beta)$, where

$$f_i^N(\beta) = a_i^N (\frac{\eta}{\eta+\Delta})^\beta - b_i^N, \quad a_i^N = \mu_2 \sum_{t=0}^{N-1} \sum_{j=2}^{k} P_t^B(i,j)$$

and $b_i^N = \mu_3 \sum_{t=0}^{N-1} \sum_{j=3}^{k} P_t^T(i,j)$. By the induction on $N \ge 4$, a_i^N and b_i^N are strictly positive for all $0 \le i \le k$. Note that $f_i^N(\beta)$ is monotonically decreasing in $\mathbb{R}$ and $f_i^N(\beta) = 0$, when it holds that

$$\beta = \beta_i^{BT}(N) = \log_{\frac{\eta}{\eta+\Delta}}(b_i^N / a_i^N).$$

Hence, we must have $V_N^B(i) > V_N^T(i)$ when $\beta < \beta_i^{BT}(N)$, and $V_N^B(i) < V_N^T(i)$ when $\beta > \beta_i^{BT}(N)$. If there exists constants $c_B, c_T \in (0,1]$ such that $\sum_{j=2}^k P_t^B(i,j) \ge c_B$ and $\sum_{j=3}^k P_t^T(i,j) \ge c_T$ hold for every $t \ge 0$ and $0 \le i \le k$, then $\{\beta_i^{BT}(N)\}_{N=4}^\infty$ is a bounded series which must have a convergent subsequence $\{\beta_i^{BT}(N_j)\}_{j=1}^\infty$ satisfying that

$$\lim_{j\to\infty} \beta_i^{BT}(N_j) = \beta^{BT} \in [-\log_{\frac{\eta}{\eta+\Delta}}(\frac{c_B\mu_3}{\mu_2}), \log_{\frac{\eta}{\eta+\Delta}}(\frac{c_T\mu_3}{\mu_2})].$$

Since bipartite and tripartite policies induce irreducible Markov chains with finite states, then $g^B = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^B(i)$

[4] Due to page limit, we can only present the proof sketch.

and $g^T = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^T(i)$ hold. Therefore, when $\beta < \beta^{BT}$, we have $g^B > g^T$, while when $\beta > \beta^{BT}$, we have $g^B < g^T$.

Remark: Theorem 1 explains how the weights w_B and w_T affect the weighted throughput of bipartite and tripartite policies based on a weight parameter β in *a family of weight functions*. Specifically, it shows that, irrespective of the presence or absence of qubit decoherence, when the weights w_B and w_T are exponential functions, e.g., $w_B = 2^\beta$ and $w_T = 3^\beta$, with a weight parameter β, as long as the transition probabilities of the bipartite and tripartite policies have non-zero lower bounds c_B and c_T, respectively, one can find a threshold β^{BT} such that the tripartite policy yields a superior weighted throughput than the bipartite policy when β is greater than the threshold β^{BT}, while the bipartite policy outperforms the tripartite policy in the weighted throughput when β falls below β^{BT}.

B. Mixed Policy vs. Bipartite and Tripartite Policies

While we restrict to bipartite or tripartite policy in Section III-A, here we further extend our model to allow the possibility that the quantum switch can probabilistically select between the bipartite action $\pi^B(S_i)$ and the tripartite action $\pi^T(S_i)$ in each state S_i.

We first introduce a new (mixed) action, denoted as A_4, into our action space $\mathcal{A}$. This action allows the quantum switch to choose $\pi^B(S_i)$ for a fraction f_i of the time and $\pi^T(S_i)$ for the remaining fraction $1 - f_i$ for state S_i, where $f_i \in (0, 1), 0 \le i \le k$. We allow the mix parameters f_i varying across states to enhance the generalizability of our framework. Since A_4 contributes differently to the throughput in various states due to varying mix parameters f_i in state S_i, $0 \le i \le k$, we assign a state-dependent weight to A_4. Specifically, the reward of action A_4 is defined as $r(S_i, A_4) = w_{M,2}(f_2\mu_2)1\{i = 2\} + w_{M,i}(f_i\mu_2 + (1-f_i)\mu_3)1\{3 \le i \le k\}$, where $w_{M,i} > 0$ represents the weight assigned to the mixed-type user entanglement for state S_i[5]. Under this extended setting with the action space $\mathcal{A} = \{A_0, A_2, A_3, A_4\}$, the expected average reward can be written as

$$g^\pi = \vec{w}_B^\pi \cdot \vec{C}_B^\pi + \vec{w}_T^\pi \cdot \vec{C}_T^\pi, \tag{6}$$

where $\vec{w}_B^\pi$ and $\vec{w}_T^\pi$ are state-dependent weight vectors satisfying $\vec{w}_B^\pi(i) = w_B 1\{\pi(S_i) = A_2\} + w_{M,i} f_i 1\{\pi(S_i) = A_4\}$ and $\vec{w}_T^\pi(i) = w_T 1\{\pi(S_i) = A_3\} + w_{M,i}(1 - f_i)1\{\pi(S_i) = A_4\}$, while $\vec{C}_B^\pi$ and $\vec{C}_T^\pi$ are bipartite and tripartite entanglement throughput vectors, satisfying $\vec{C}_B^\pi(i) = \Pi(i)\mu_2$ and $\vec{C}_T^\pi(i) = \Pi(i)\mu_3$, which are the bipartite and tripartite entanglement throughput in state S_i under the policy π, $0 \le i \le k$, respectively. Compared with Eq.4, Eq.6 serves as a more general expression of the weighted throughput of a quantum switch under the policy π. This generalization is achieved by assigning state-dependent weights to the bipartite and tripartite entanglement throughput, offering greater flexibility for customizations. Note that when $w_B = 1$, $w_T = 1$ and $w_{M,i} = 1, 2 \le i \le k$, g^π represents the *total throughput* of the quantum switch. Our optimization goal in this extended setting remains to determine the optimal policy π that maximizes the weighted throughput g^π defined in Eq.6.

Based on the definition of the mixed action A_4, the mixed policy π^M is defined as $\pi^M(S_i) = A_0 1\{0 \le i \le 1\} + A_4 1\{2 \le i \le k\} = A_0 1\{0 \le i \le 1\} + (f_2 A_2 + (1 - f_2)A_0)1\{i = 2\} + (f_i A_2 + (1-f_i)A_3)1\{3 \le i \le k\}$. A specific mix vector $\vec{f} = (f_0, \ldots, f_k)$ determines a unique mixed policy π^M. The reward for π^M is given by $r^M(i) = w_{M,2} f_2 \mu_2 1\{i = 2\} + w_{M,i}(f_i\mu_2 + (1 - f_i)\mu_3)1\{3 \le i \le k\}$. The transition probability matrix P^M for the Markov chain under π^M, depicted in Figure 4, is derived through uniformization of the transition rate matrix Q^M, with a uniform rate ρ. The matrix Q^M is constructed to satisfy $Q^M(i, j) = f_i Q^B(i, j) + (1 - f_i)Q^T(i, j)$ for all $0 \le i, j \le k$. Also, the total expected reward $V_N^{\pi^M}(i)$ for state $S_i, 0 \le i \le k$, under π^M can be expressed as $V_N^M(i)$ adhering to Eq.1. Since any mixed policy with the mix vector $\vec{f}$ always induces an irreducible Markov chain with finite states, the expected average reward $g^M = g^{\pi^M}$ must exist.

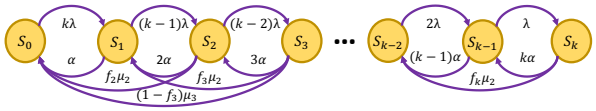


Fig. 4: The transition rate diagram of the Markov chain determined by the mixed policy π^M.

In order to simulate different relationships between bipartite, tripartite and mixed-type user entanglement, we consider the weights of mixed-mode basis measurement to be a function of both η and Δ, similar to w_B and w_T, with a weight parameter $\beta \in (-\infty, +\infty)$, defined as $w_{M,i} = (f_i\eta + (1 - f_i)(\eta + \Delta))^\beta$ for all $2 \le i \le k$. However, $w_{M,i}$ incorporates the mix parameter f_i, allowing the weight to vary depending on the specific mixed-mode basis measurement being considered.

For each unique mixed policy, we carry out a comparative analysis with both the bipartite and tripartite policies, irrespective of whether the stored qubits go through decoherence or not. Let us first compare the mixed policy against the bipartite policy, Theorem 2 states that when the weight parameters $w_{M,i}$ of the mixed-mode basis measurement follow exponential functions, with certain constraints on the transition probability matrix P^B and P^M, a mixed policy, defined by a vector $\vec{f}$, is superior to a bipartite policy in the weighted throughput when the weight parameter β is greater than a threshold β^{BM}, while the situation reverses if β is smaller than the threshold β^{BM}. Therefore, one can conclude that incorporating the tripartite policy into a mixed strategy enables the mixed policy becomes advantageous in achieving higher weighted throughput as β increases, while the straightforward approach of the bipartite policy is preferable in the scenario when one put higher weight on bipartite user entanglement.

Theorem 2. *Suppose* $w_B = \eta^\beta$, $w_{M,i} = (f_i\eta + (1 - f_i)(\eta + \Delta))^\beta$ *for* $2 \le i \le k$, *where* $\eta > 1$ *and* $\Delta > 0$ *are*

[5]Note that our use of a different weight parameter $w_{M,i}$ represents a significant departure from the use of mixed policies in classical MDP theory, and thus existing results on the optimality of deterministic vs. mixed policies [36] do not directly apply.

constants, $\rho > \max\{\max_{0\le i\le k}|Q^B(i,i)|, \max_{0\le i\le k}|Q^M(i,i)|\}$ *and decoherence rate* $\alpha \ge 0$*. If there exists a constant* $c_{BM} > 0$ *such that*

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} P_t^M(i,j) f_j (\frac{\eta + (1-f_j)\Delta}{\eta})^{c_{BM}} \ge N$$

and

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} (P_t^B(i,j) - P_t^M(i,j)) > 0$$

hold for $0 \le i \le k$ *and for all* $N \ge 3$*. Then there must exist a threshold* $\beta^{BM} = \beta^{BM}(\vec{f}) \in [0, c_{BM}]$ *such that* $g^B > g^M$ *if* $\beta < \beta^{BM}$*, and* $g^B < g^M$ *if* $\beta > \beta^{BM}$*.*

Proof sketch. For any $N \ge 0$, we can deduce that $V_N^B(i) - V_N^M(i) = \eta^\beta f_i^N(\beta)$, where

$$f_i^N(\beta) = a_i^N - \sum_{j=2}^{k} c_{i,j}^N (\frac{\eta + (1-f_j)\Delta}{\eta})^\beta,$$

$c_{i,2}^N = \sum_{t=0}^{N-1} P_t^M(i,2) f_2\mu_2$, $c_{i,j}^N = \sum_{t=0}^{N-1} P_t^M(i,j)(f_j\mu_2 + (1-f_j)\mu_3), 3 \le j \le k, a_i^N = \mu_2 \sum_{t=0}^{N-1}\sum_{j=2}^{k} P_t^B(i,j)$. Specifically, we can know that a_i^N and $c_{i,j}^N$ are strictly positive for all $2 \le j \le k$ and $0 \le i \le k$ by the induction on $N \ge 3$. If there exists a constant $c_{BM} > 0$ satisfies that

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} P_t^M(i,j) f_j (\frac{\eta + (1-f_j)\Delta}{\eta})^{c_{BM}} \ge N$$

and

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} (P_t^B(s,j) - P_t^M(i,j)) > 0$$

hold for $0 \le i \le k$ and for all $N \ge 3$, then $f_i^N(0) > 0$ and $f_i^N(c_{BM}) < 0$. Since $f_i^N(\beta)$ is a monotonically decreasing function in $\mathbb{R}$, there must exists a threshold $\beta_i^{BM}(N) \in [0, c_{BM}]$ such that $V_N^B(i) > V_N^M(i)$ when $\beta < \beta_i^{BM}(N)$ and $V_N^B(i) < V_N^M(i)$ when $\beta > \beta_i^{BM}(N)$. For any $0 \le i \le k$, $\{\beta_i^{BM}(N)\}_{N=3}^\infty$ is a bounded series, so it must have a convergent subsequence $\{\beta_i^{BM}(N_j)\}_{j=1}^\infty$ satisfying that $\lim_{j\to\infty} \beta_i^{BM}(N_j) = \beta^{BM} \in [0, c_{BM}]$. Since bipartite and mixed policies induce irreducible Markov chains with finite states, then $g^B = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^B(i)$ and $g^M = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^M(i)$ hold. Then we have $g^B > g^M$ when $\beta < \beta^{BM}$ and $g^B < g^M$ when $\beta > \beta^{BM}$.

Furthermore, we can compare the mixed policy with the tripartite policy. Theorem 3 states that under certain constraints on the transition probabilities of the mixed and tripartite policies, the weighted throughput of a specific mixed policy will be inferior to a tripartite policy if the weight parameter β surpasses a threshold β^{MT}. This reflects the increasing weighted throughput of exclusively using tripartite basis measurements as β increases. Conversely, if $\beta < \beta^{MT}$, the mixed policy is superior to the tripartite policy, suggesting the value of incorporating bipartite basis measurements.

Theorem 3. *Suppose* $w_T = (\eta + \Delta)^\beta$, $w_{M,i} = (f_i\eta + (1-f_i)(\eta+\Delta))^\beta$ *for* $2 \le i \le k$*, where* $\eta > 1$ *and* $\Delta > 0$ *are constants,* $\rho > \max\{\max_{0\le i\le k}|Q^M(i,i)|, \max_{0\le i\le k}|Q^T(i,i)|\}$ *and decoherence rate* $\alpha \ge 0$*. If there exists a constant* $c_{MT} > 0$ *such that*

$$\sum_{t=0}^{N-1}\sum_{j=3}^{k} P_t^T(i,j) \ge N\frac{\mu_2}{\mu_3} \max_{2\le j\le k}\{(\frac{\eta + (1-f_j)\Delta}{\eta+\Delta})^{c_{MT}}\}$$

and

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} P_t^M(i,j) f_j\mu_2 \ge N\mu_3$$

hold for $0 \le i \le k$ *and for all* $N \ge 4$*, then there exists a threshold* $\beta^{MT} = \beta^{MT}(\vec{f}) \in [0, c_{MT}]$ *such that* $g^M > g^T$ *when* $\beta < \beta^{MT}$*, and* $g^M < g^T$ *when* $\beta > \beta^{MT}$*.*

Proof sketch. For any $N \ge 0$, we can deduce that $V_N^M(i) - V_N^T(i) = (\eta+\Delta)^\beta f_i^N(\beta)$, where

$$f_i^N(\beta) = \sum_{j=2}^{k} c_{i,j}^N (\frac{\eta + (1-f_j)\Delta}{\eta+\Delta})^\beta - b_i^N,$$

$c_{i,2}^N = \sum_{t=0}^{N-1} P_t^M(i,2) f_2\mu_2$, $c_{i,j}^N = \sum_{t=0}^{N-1} P_t^M(i,j)(f_j\mu_2 + (1-f_j)\mu_3), 3 \le j \le k$, and $b_i^N = \sum_{t=0}^{N-1}\sum_{j=3}^{k} P_t^T(i,j)\mu_3$. Specifically, we can know that $c_{i,j}^N > 0$ and $b_i^N > 0$ for all $2 \le j \le k$ and $0 \le i \le k$ by the induction on $N \ge 4$. If there exists a constant $c_{MT} > 0$ such that

$$\sum_{t=0}^{N-1}\sum_{j=3}^{k} P_t^T(i,j) \ge N\frac{\mu_2}{\mu_3} \max_{2\le j\le k}\{(\frac{\eta + (1-f_j)\Delta}{\eta+\Delta})^{c_{MT}}\}$$

and

$$\sum_{t=0}^{N-1}\sum_{j=2}^{k} P_t^M(i,j) f_j\mu_2 \ge N\mu_3$$

hold for $0 \le i \le k$ and all $N \ge 4$, then we have $f_i^N(0) > 0$ and $f_i^N(c_{MT}) < 0$. Since $f_i^N(\beta)$ is a monotonically decreasing function in $\mathbb{R}$, there must exist a threshold $\beta_i^{MT}(N) \in [0, c_{MT}]$ such that $V_N^M(i) > V_N^T(i)$ when $\beta < \beta_i^{MT}(N)$ and $V_N^M(i) < V_N^T(i)$ when $\beta > \beta_i^{MT}(N)$. Note that, for any $0 \le i \le k$, $\{\beta_i^{MT}(N)\}_{N=4}^\infty$ is a bounded series. Then it must have a convergent subsequence $\{\beta_i^{MT}(N_j)\}_{j=1}^\infty$ satisfying that $\lim_{j\to\infty} \beta_i^{MT}(N_j) = \beta^{MT} \in [0, c_{MT}]$. Since tripartite and mixed policies induce irreducible Markov chains with finite states, then $g^T = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^T(i)$ and $g^M = \lim_{j\to\infty} \frac{1}{N_j} V_{N_j}^M(i)$ hold. Hence, we have $g^M < g^T$ when $\beta > \beta^{MT}$ and $g^M > g^T$ when $\beta < \beta^{MT}$.

Remark: Different from Theorem 1, which contrasts two specific policies, Theorem 2 and 3 broaden the analysis to compare bipartite and tripartite policies against *an infinite spectrum of mixed policies*, each specified by a vector $\vec{f}$. Collectively, Theorems 1, 2, and 3 imply that the optimal entanglement policy of the quantum switch can be easily adapted in response to the weight configuration of the system. These insights enable decision-makers to make simple decisions on actions based on current dynamics, thereby enhancing network throughput.

IV. Numerical Experiments

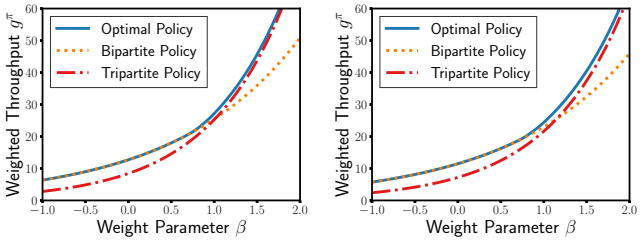
In this section, we conduct experiments to substantiate our theoretical analysis on the dynamics between the bipartite and tripartite policies, as well as the interplay between bipartite, tripartite and mixed policies in the presence and/or absence of decoherence. In these experiments, we obtain the optimal entanglement policy by employing the relative value iteration algorithm [36] within the MDP framework[6], with the aim of maximizing the weighted throughput. Then, we show that the optimal entanglement policy exhibits a threshold-based and state-dependent property.

Experimental Setup: We consider a quantum switch serving $k = 8$ users with a buffer size $B = 1$. For a stable system, the new link-level entanglement arrival rate is lower than the basis measurement rates [45]. Besides, the quantum switch more easily performs bipartite basis measurement than tripartite basis measurement [37], i.e., $\mu_3 < \mu_2$. We set $\lambda = 6$ as the link-level entanglement arrival rate for each link. The mean times to successfully perform bipartite and tripartite basis measurements are $1/\mu_2 = 1/15$ and $1/\mu_3 = 1/12$. For the weight functions, we set $w_B = 2^\beta$, $w_T = 3^\beta$ and $w_{M,i} = (3 - f_i)^\beta$ for $2 \leq i \leq k$. If decoherence exists, the decoherence rate is set to $\alpha = 2.5$. The fraction vector in the mixed policy is set as $\vec{f} = (1, 1, 0.95, 0.78, 0.25, 0.96, 0.35, 0.65, 0.78)$. We use a convergence threshold of $\epsilon = 10^{-3}$ in the relative value iteration algorithm.

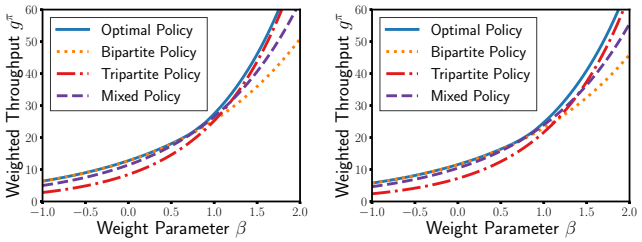
Experiment 1: Existence of thresholds between bipartite, tripartite and mixed policies: We intend to reveal (a) the existence of thresholds β^{BT}, β^{BM} and β^{MT} in Theorem 1, 2 and 3, respectively, and (b) the relationship between the optimal policy and the bipartite, mixed and tripartite policies.

For bipartite, tripartite and mixed policies, we obtain the expected average rewards g^B, g^T and g^M under different β. Here, we consider the range of β in $(-1, 2)$. We also obtain g^*, which is the expected average reward of the optimal policy. As shown in Figure 5(a), the curves of the g^B (in orange) and g^T (in red) have a distinct intersection at $\beta^{BT} \approx 1$. When $\beta < \beta^{BT}$, the expected average reward g^B of the bipartite policy is higher than that of the tripartite policy g^T, while the situation reverses when $\beta > \beta^{BT}$, which validates Theorem 1. When $\beta < \beta^{BT}$, the expected average reward of the optimal policy g^* is closer to the bipartite policy g^B than the tripartite policy g^T. Therefore, bipartite action $\pi^B(s)$ will be applied to most states as the optimal action in this range of β. We validate this in Experiment 2, where we present the optimal state-dependent policies in TABLE I. In the first 4 rows in TABLE I, when $\beta \in (-5, 0.753)$, the optimal actions for most states are A_2. Similarly, when $\beta > \beta^{BT}$, the value of g^T is closer to that of g^*, indicating tripartite action $\pi^T(s)$ is applied to most states as the optimal action. From TABLE I, we see that as β increases, A_3 will gradually become the dominant action, and the optimal entanglement policy becomes the tripartite policy when $\beta > 2.413$.

[6]For the implementation of the relative value iteration algorithm, we refer to the code in https://pymdptoolbox.readthedocs.io/en/latest/.



(a) $\lambda = 6, \alpha = 0$ (b) $\lambda = 6, \alpha = 2.5$



(c) $\lambda = 6, \alpha = 0$ (d) $\lambda = 6, \alpha = 2.5$

Fig. 5: The expected average reward g^π with different values of weight parameter β under different policies.

We also incorporate a mixed policy into the MDP framework and obtain the values of g^M under different values of β as depicted in Figure 5(c). First, we observe that the expected rewards of the tripartite policy g^T (in red) and the mixed policy g^M (in purple) intersect when $\beta = \beta^{MT} \approx 1.2 \in (1, 1.5)$. In particular, when $\beta < \beta^{MT}$, the expected reward of the mixed policy g^M is higher than that of the tripartite policy g^M, and the situation changes when $\beta > \beta^{MT}$, which validates Theorem 3. Similarly, we observe the intersection of the expected reward of the bipartite policy (in orange) and the mixed policy (in purple) when $\beta = \beta^{BM} \approx 0.8 \in (0.5, 1)$ in Figure 5(c), which validates Theorem 2. Moreover, one can see $\beta^{BM} < \beta^{BT} < \beta^{MT}$, the expected reward of the mixed policy and the bipartite policy are closer to the optimal policy when β is small, which indicates that the actions of the mixed policy and bipartite policy are more likely to be the optimal actions in this range of β. However, as β increases, the expected reward of the tripartite policy is closer to that of the optimal policy than the other two policies, indicating that most actions in the optimal policy are tripartite actions. TABLE III provides support for our claims, which shows that A_2 and A_4 are dominant actions when β is small, while A_3 gradually replaces A_2 and A_4 as β increases and finally, the optimal policy becomes the tripartite policy when $\beta > 2.678$.

Next, let us examine the effect of the decoherence on the threshold-based entanglement policy. We notice that the existence of decoherence will have a negative effect on taking tripartite action as the optimal action. Indeed, as α increases, the quantum switch will adopt a conservative strategy and prefer the bipartite action. Specifically, in Figure 5(a) and 5(b), one can observe that the value of β^{BT} corresponding to the intersection of the curves of bipartite policy and tripartite policy under $\alpha = 2.5$ is higher than that of $\alpha = 0$. This reveals that under decoherence, it requires higher value of β to allow the A_3 to be more advantageous than the A_2. Since

Range of β	Optimal Policy								
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$(-5.000, 0.673)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_2
$(0.674, 0.681)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_3
$(0.682, 0.747)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_3	A_3
$(0.748, 0.753)$	A_0	A_0	A_2	A_2	A_2	A_3	A_2	A_3	A_3
$(0.754, 0.823)$	A_0	A_0	A_2	A_2	A_2	A_3	A_3	A_3	A_3
$(0.824, 0.955)$	A_0	A_0	A_2	A_3	A_2	A_3	A_3	A_3	A_3
$(0.956, 2.412)$	A_0	A_0	A_2	A_3	A_3	A_3	A_3	A_3	A_3
$(2.413, 5.000)$	A_0	A_0	A_0	A_3	A_3	A_3	A_3	A_3	A_3

TABLE I: Optimal policies in different ranges of β when the decoherence rate is $\alpha = 0$.

Range of β	Optimal Policy								
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$(-5.000, 0.661)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_2
$(0.662, 0.676)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_3
$(0.677, 0.734)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_3	A_3
$(0.735, 0.735)$	A_0	A_0	A_2	A_2	A_2	A_2	A_3	A_3	A_3
$(0.736, 0.789)$	A_0	A_0	A_2	A_2	A_2	A_3	A_3	A_3	A_3
$(0.790, 0.941)$	A_0	A_0	A_2	A_3	A_2	A_3	A_3	A_3	A_3
$(0.942, 2.731)$	A_0	A_0	A_2	A_3	A_3	A_3	A_3	A_3	A_3
$(2.732, 5.000)$	A_0	A_0	A_0	A_3	A_3	A_3	A_3	A_3	A_3

TABLE II: Optimal policies in different ranges of β when the decoherence rate is $\alpha = 2.5$.

performing tripartite basis measurement takes longer time than bipartite basis measurement ($\mu_2 > \mu_3$), it is not profitable for the quantum switch to adopt tripartite policy. TABLE II also justifies that with decoherence, the quantum switch prefers A_2 unless the weight parameter β is sufficiently large. Similar phenomenon also holds when the mixed policy is involved, by comparing Figure 5(c) and Figure 5(d), which is supported by the comparison between TABLE III and TABLE IV.

Experiment 2. Existence of a threshold-based optimal policy and its dynamics: We will further examine the optimal entanglement policy derived from the relative value iteration algorithm, and empirically reveal its structure. Specifically, we aim to (1) reveal that the optimal entanglement policy changes from the bipartite policy π^B to tripartite policy π^T as β increases from a purely bipartite region to a purely tripartite region, and to identify the optimal entanglement policies between two regions; (2) show that, even in the range where π^B and π^T are not optimal, the actions of the optimal entanglement policy at each state S_i will change from the bipartite action $\pi^B(S_i)$ to tripartite action $\pi^T(S_i)$ as the weight parameter β exceeds an empirical threshold $\tilde{\beta}_{S_i}^{BT}$.

In this experiment, we consider a larger range of weight parameter $\beta \in (-5, 5)$. TABLE I and II present the optimal policies when decoherence exists or not. Observe that TABLE I and II are consistent with Theorem 1 because as β becomes small or large, the optimal entanglement policy eventually becomes the bipartite policy or the tripartite policy, whether decoherence exists or not. We denote the region where the optimal entanglement policy is the tripartite policy for all $\beta \geq \beta_T^*$ as the "*purely tripartite region*". Similarly, we denote the region where the optimal policy is the bipartite policy for all $\beta \leq \beta_B^*$ as the "*purely bipartite region*". In TABLE I, the interval $(2.413, 5)$ is the "*purely tripartite region*" (highlighted in red). Besides, the "*purely bipartite region*" exists when $\beta \in (-5, 0.673)$ (highlighted in green). Knowing this threshold structure benefits the decision-making. For instance, for a given β', we can obtain the optimal entanglement policy through the relative value iteration algorithm [36]. If the optimal entanglement policy is the bipartite policy, it remains so for $\beta < \beta'$; if the tripartite policy is optimal, it also holds for $\beta > \beta'$. For the weight parameters between β_B^* and β_T^*, we can use the relative value iteration algorithm to efficiently determine the optimal entanglement policy.

The above two tables also shows that even for the range of β where neither π^B nor π^T are optimal, there is a tendency for the optimal entanglement action of a given state S_j to shift from bipartite action to tripartite action as β increases. For instance, the optimal entanglement action for state S_3 (highlighted in blue) in TABLE I transits from A_2 to A_3 as β evolves from $(-5, 0.823)$ to $(0.824, 5)$. Hence, there exists a threshold $\tilde{\beta}_{S_3}^{BT}$ for state S_3 such that the tripartite action will be used at state S_3 for $\beta \geq \tilde{\beta}_{S_3}^{BT}$. From TABLE I, such a threshold for state S_3 lies in $(0.823, 0.824)$. Note that the thresholds $\tilde{\beta}_{S_j}^{BT}$ with respect to different states S_j may be different, which creates a gap between the *purely tripartite region* and the *purely bipartite region*. This new insight reveals that the shift in the optimal entanglement action for each state S_i from $\pi^B(S_i)$ to $\pi^T(S_i)$ triggers the transition of the optimal entanglement policy from π^B to π^T.

Experiment 3. Existence of Threshold-based Optimal Policy under the Mixed Policy Setting: To further explore optimal entanglement policies of the quantum switch, we include a mixed policy, introducing an additional action A_4. Utilizing the relative value iteration algorithm, we determine the optimal entanglement policy across a spectrum of β values within $(-5, 5)$. The findings, as presented in TABLE III and TABLE IV, show the optimal actions in the absence or presence of decoherence, respectively. Recall that Theorem 2 and 3 formally specify the dynamics between the mixed, bipartite, and tripartite policies. In particular, a threshold phenomenon is revealed that governs their interplay, irrespective of the presence or absence of decoherence. TABLE III offers an illustration of this dynamic, indicating that there is a progressive shift in the quantum switch's optimal entanglement policy: from a bipartite to a mixed, and eventually, to a tripartite policy, as the weight parameter β increases. For instance, state S_6 initially favors action A_2, and transits to A_4 within the β interval of $(0.7, 0.8)$, and finally settling on action A_3 as β further increases. From TABLE III, when β is sufficiently large, one can observe the existence of a purely tripartite region $(2.7, 5)$, indicating that the quantum switch should follow tripartite policy if the weight of A_3 is much greater than the weights of A_4 and A_2 . One can also observe the existence of a purely bipartite region $(-5, -0.8)$ when one put a relatively higher weight on A_2 than that of A_3 and A_4.

Similar transitions of optimal entanglement actions under the decoherence effect are shown in TABLE IV.

Range of β	Optimal Policy								
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$(-5.000, -0.800)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_2
$(-0.700, 0.500)$	A_0	A_0	A_4	A_2	A_2	A_2	A_2	A_2	A_2
$(0.600, 0.600)$	A_0	A_0	A_4	A_2	A_2	A_4	A_2	A_4	A_4
$(0.700, 0.700)$	A_0	A_0	A_4	A_4	A_2	A_4	A_4	A_4	A_4
$(0.800, 0.800)$	A_0	A_0	A_4	A_4	A_2	A_3	A_4	A_3	A_3
$(0.900, 0.900)$	A_0	A_0	A_4	A_3	A_2	A_3	A_3	A_3	A_3
$(1.000, 1.100)$	A_0	A_0	A_4	A_3	A_4	A_3	A_3	A_3	A_3
$(1.200, 2.600)$	A_0	A_0	A_4	A_3	A_3	A_3	A_3	A_3	A_3
$(2.700, 5.000)$	A_0	A_0	A_0	A_3	A_3	A_3	A_3	A_3	A_3

TABLE III: Optimal policies in different ranges of β when $\alpha = 0$ after incorporating the mixed policy into MDP.

Range of β	Optimal Policy								
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$(-5.000, -0.700)$	A_0	A_0	A_2	A_2	A_2	A_2	A_2	A_2	A_2
$(-0.600, 0.500)$	A_0	A_0	A_4	A_2	A_2	A_2	A_2	A_2	A_2
$(0.600, 0.600)$	A_0	A_0	A_4	A_2	A_2	A_4	A_2	A_4	A_4
$(0.700, 0.700)$	A_0	A_0	A_4	A_4	A_2	A_4	A_4	A_4	A_4
$(0.800, 0.800)$	A_0	A_0	A_4	A_4	A_2	A_3	A_4	A_3	A_3
$(0.900, 1.100)$	A_0	A_0	A_4	A_3	A_4	A_3	A_3	A_3	A_3
$(1.200, 3.000)$	A_0	A_0	A_4	A_3	A_3	A_3	A_3	A_3	A_3
$(3.100, 5.000)$	A_0	A_0	A_0	A_3	A_3	A_3	A_3	A_3	A_3

TABLE IV: Optimal policies in different ranges of β when $\alpha = 2.5$ after incorporating the mixed policy into MDP.

V. Conclusion

In this work, we explore the optimal entanglement policies that maximize the weighted throughput of a quantum switch by examining its structural properties based on existing link-level entanglements. We use an MDP framework to model the dynamics of the quantum switch and theoretically prove the equivalence between the expected average reward and the weighted throughput. Based on the MDP formulation, we theoretically unveil the threshold-based relationships among three special entanglement policies: bipartite policy, tripartite policy, and mixed policy within a family of weight functions with a weight parameter β. Our theoretical analysis establishes that the tripartite policy outperforms the bipartite policy in weighted throughput when the weight parameter β exceeds a threshold β^{BT}. Conversely, the bipartite policy is superior to the tripartite policy when β falls below β^{BT}. We extend this analysis to a family of mixed policies—a hybrid of bipartite and tripartite policies—and reveal a threshold-based relationship between any specific mixed policy and the bipartite/tripartite policies. Empirically, we validate our theoretical insights by showing the dynamics between the bipartite, mixed and tripartite policies, and reveal the existence of a threshold-based optimal entanglement policy. Notably, we also find that decoherence enables the quantum switch to adopt a conservative strategy to avoid the risk of losing existing link-level entanglements, preferring to perform bipartite basis measurements as soon as possible instead of waiting for the tripartite basis measurements.

Appendix

We first construct the transition rate matrix Q^a with respect to each action $a \in \mathcal{A}$, where $Q^a(i,j)$ represents the transition rate from state S_i to state S_j under action a. For action A_0, a state S_i will transit to S_{i+1} at a rate of $(k-i)\lambda$ due to the arrival of new link-level entanglement, $0 \le i \le k-1$. In addition, a state S_i can transit to state S_{i-1} at a rate of $i\alpha$ due to the decoherence effect, where $1 \le i \le k$. Hence, the transition rate matrix of action A_0 is:

$$Q^{A_0}(i,j) = \begin{cases} -\sum_{j \neq i} Q^{A_0}(i,j), & j = i, i = 0, 1 \cdots, k \\ (k-i)\lambda, & j = i+1, i = 0, \cdots, k-1 \\ i\alpha, & j = i-1, i = 1, \cdots, k \\ 0, & otherwise. \end{cases} \tag{7}$$

We derive the transition rate matrices Q^{A_2} and Q^{A_3} similarly. The only differences between Q^{A_0}, Q^{A_2} and Q^{A_3} are the additional transition rates $Q^{A_2}(i, i-2) = \mu_2, 2 \le i \le k$, and $Q^{A_3}(i, i-3) = \mu_3, 3 \le i \le k$, due to the bipartite and tripartite basis measurements, respectively. We consider an action A_4 which allows the quantum switch to probabilistically perform bipartite basis measurement and tripartite basis measurement. Specifically, when the quantum switch is in state S_i, it adopts the action from the bipartite policy $\pi^B(S_i)$ for a fraction f_i of the time, and adopts the action from the tripartite policy $\pi^T(S_i)$ for the remaining fraction $(1-f_i)$, where $0 \le i \le k$. Therefore, Q^{A_4} satisfy that $Q^{A_4}(i,j) = f_i Q^{A_2}(i,j) + (1-f_i) Q^{A_3}(i,j)$, for all $0 \le i, j \le k$. By applying the uniformization technique on each transition rate matrix, we can obtain the transition probability matrix of action a as $P^a = I + \frac{Q^a}{\rho}$, where ρ is an uniformization rate parameter and it satisfies $\rho \ge \max_{a \in \mathcal{A}} \{ \max_{0 \le i \le k} |Q^a(i,i)| \}$.

References

[1] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information*. American Association of Physics Teachers, 2002.

[2] N. Gisin and R. Thew, "Quantum communication," *Nature photonics*, vol. 1, no. 3, pp. 165–171, 2007.

[3] M. Liu, J. Allcock, K. Cai, S. Zhang, and J. C. Lui, "Quantum networks with multiple service providers: Transport layer protocols and research opportunities," *IEEE Network*, vol. 36, no. 5, pp. 56–62, 2022.

[4] M. Liu, Z. Li, K. Cai, J. Allcock, S. Zhang, and J. C. Lui, "Quantum bgp with online path selection via network benchmarking," in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications*, 2024.

[5] M. Liu, Z. Li, X. Wang, and J. C. Lui, "Linkselfie: Link selection and fidelity estimation in quantum networks," in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications*, 2024.

[6] B. He, D. Zhang, S. W. Loke, S. Lin, and L. Lu, "Building a hierarchical architecture and communication model for the quantum internet," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1919–1935, 2024.

[7] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, "Advances in quantum cryptography," *Advances in optics and photonics*, vol. 12, no. 4, pp. 1012–1236, 2020.

[8] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science," *Chemical Reviews*, vol. 120, no. 22, pp. 12 685–12 717, 2020.

[9] Y. Zhao and C. Qiao, "Distributed transport protocols for quantum data networks," *IEEE/ACM Transactions on Networking*, vol. 31, no. 6, pp. 2777–2792, 2023.

[10] Y. Zeng, J. Zhang, J. Liu, Z. Liu, and Y. Yang, "Entanglement routing design over quantum networks," *IEEE/ACM Transactions on Networking*, vol. 32, no. 1, pp. 352–367, 2023.

[11] ——, "Multi-entanglement routing design over quantum networks," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 510–519.

[12] L. Chen, K. Xue, J. Li, Z. Li, R. Li, N. Yu, Q. Sun, and J. Lu, "Redp: Reliable entanglement distribution protocol design for large-scale quantum networks," *IEEE Journal on Selected Areas in Communications*, 2024.

[13] J.-L. Jiang, M.-X. Luo, and S.-Y. Ma, "Quantum network capacity of entangled quantum internet," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1900–1918, 2024.

[14] S. Pouryousef, N. K. Panigrahy, and D. Towsley, "A quantum overlay network for efficient entanglement distribution," in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[15] Y. Mao, Y. Liu, and Y. Yang, "Qubit allocation for distributed quantum computing," in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[16] L. Yang, Y. Zhao, L. Huang, and C. Qiao, "Asynchronous entanglement provisioning and routing for distributed quantum computing," in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[17] C. Sekga and M. Mafu, "Tripartite quantum key distribution implemented with imperfect sources," *Optics*, vol. 3, no. 3, pp. 191–208, 2022.

[18] P. W. Shor and J. Preskill, "Simple proof of security of the bb84 quantum key distribution protocol," *Physical review letters*, vol. 85, no. 2, p. 441, 2000.

[19] M. S. Akhtar, G. Krishnakumar, B. Vishnu, and A. Sinha, "Fast and secure routing algorithms for quantum key distribution networks," *IEEE/ACM Transactions on Networking*, 2023.

[20] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, pp. 575–579, 1997.

[21] P. Nain, G. Vardoyan, S. Guha, and D. Towsley, "On the analysis of a multipartite entanglement distribution switch," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2020.

[22] ——, "Analysis of a tripartite entanglement distribution switch," *Queueing Systems*, vol. 101, no. 3-4, pp. 291–328, 2022.

[23] G. Vardoyan, S. Guha, P. Nain, and D. Towsley, "On the stochastic analysis of a quantum entanglement switch," *ACM SIGMETRICS Performance Evaluation Review*, vol. 47, no. 2, pp. 27–29, 2019.

[24] G. Vardoyan, P. Nain, S. Guha, and D. Towsley, "On the capacity region of bipartite and tripartite entanglement switching," *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, pp. 1–16, 2019.

[25] W. Dai, A. Rinaldi, and D. Towsley, "The capacity region of entanglement switching: Stability and zero latency," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2022, pp. 389–399.

[26] T. Vasantam and D. Towsley, "A throughput optimal scheduling policy for a quantum switch," in *Quantum Computing, Communication, and Simulation II*, vol. 12015. SPIE, 2022, pp. 14–23.

[27] V. Valls, P. Promponas, and L. Tassiulas, "On the capacity of the quantum switch with and without entanglement decoherence," *IEEE Communications Letters*, vol. 27, no. 9, pp. 2388–2392, 2023.

[28] P. Promponas, V. Valls, S. Guha, and L. Tassiulas, "Maximizing entanglement rates via efficient memory management in flexible quantum switches," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1749–1762, 2024.

[29] J. Huang and L. Huang, "Learning-based optimal quantum switch scheduling," *SIGMETRICS Perform. Eval. Rev.*, vol. 51, no. 2, pp. 75–77, oct 2023.

[30] N. K. Panigrahy, T. Vasantam, D. Towsley, and L. Tassiulas, "On the capacity region of a quantum switch with entanglement purification," in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[31] I. Herbauts, B. Blauensteiner, A. Poppe, T. Jennewein, and H. Hübel, "Demonstration of active routing of entanglement in a multi-user network," *Opt. Express*, vol. 21, no. 23, pp. 29 013–29 024, Nov 2013.

[32] M. M. Cunha, A. Fonseca, and E. O. Silva, "Tripartite entanglement: Foundations and applications," *Universe*, vol. 5, no. 10, 2019. [Online]. Available: https://www.mdpi.com/2218-1997/5/10/209

[33] W. McCutcheon, A. Pappa, B. A. Bell, A. Mcmillan, A. Chailloux, T. Lawson, M. Mafu, D. Markham, E. Diamanti, I. Kerenidis *et al.*, "Experimental verification of multipartite entanglement in quantum networks," *Nature communications*, vol. 7, no. 1, p. 13251, 2016.

[34] L. Benmohamed and S. Meerkov, "Feedback control of congestion in packet switching networks: the case of a single congested node," *IEEE/ACM Transactions on Networking*, vol. 1, no. 6, pp. 693–708, 1993.

[35] G. Vardoyan, S. Guha, P. Nain, and D. Towsley, "On the exact analysis of an idealized quantum switch," *ACM SIGMETRICS Performance Evaluation Review*, vol. 48, no. 3, pp. 79–80, 2021.

[36] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 4.

[37] Z. Su, "Generating tripartite nonlocality from bipartite resources," *Quantum Information Processing*, vol. 16, p. 28, 12 2016.

[38] K. F. Reim, P. Michelberger, K. C. Lee, J. Nunn, N. K. Langford, and I. A. Walmsley, "Single-photon-level quantum memory at room temperature," *Phys. Rev. Lett.*, vol. 107, p. 053603, Jul 2011. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.107.053603

[39] M. L. Puterman, "Markov decision processes," *Handbooks in operations research and management science*, vol. 2, pp. 331–434, 1990.

[40] S. Khatri, "On the design and analysis of near-term quantum network protocols using markov decision processes," *AVS Quantum Science*, vol. 4, no. 3, 2022.

[41] Á. G. Iñesta, G. Vardoyan, L. Scavuzzo, and S. Wehner, "Optimal entanglement distribution policies in homogeneous repeater chains with cutoffs," *npj Quantum Information*, vol. 9, no. 1, p. 46, 2023.

[42] V. Kumar, N. K. Chandra, K. P. Seshadreesan, A. Scheller-Wolf, and S. Tayur, "Optimal entanglement distillation policies for quantum switches," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 1. IEEE, 2023, pp. 1198–1204.

[43] W. J. Stewart, *Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling*. Princeton University Press, 2009.

[44] O. C. Ibe, *Markov Processes for Stochastic Modeling*. Academic Press, 2009.

[45] T. Coopmans, R. Knegjens, A. Dahlberg, D. Maier, L. Nijsten, J. de Oliveira Filho, M. Papendrecht, J. Rabbie, F. Rozpedek, M. Skrzypczyk *et al.*, "Netsquid, a network simulator for quantum information using discrete events," *Communications Physics*, vol. 4, no. 1, p. 164, 2021.