On the Capacity Region of a Quantum Switch with Entanglement Purification

Nitish K. Panigrahy*, Thirupathaiah Vasantam †, Don Towsley ‡ and Leandros Tassiulas*
*Yale University, USA. †Durham University, UK. ‡University of Massachusetts Amherst, USA.
Email: *{nitishkumar.panigrahy, leandros.tassiulas}@yale.edu, †thirupathaiah.vasantam@durham.ac.uk, ‡towsley@cs.umass.edu

IEEE INFOCOM 2023 - IEEE Conference on Computer Communications | 979-8-3503-3414-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/INFOCOM53939.2023.10229003

***Abstract*—Quantum switches are envisioned to be an integral component of future entanglement distribution networks. They can provide high quality entanglement distribution service to end-users by performing quantum operations such as entanglement swapping and entanglement purification. In this work, we characterize the capacity region of such a quantum switch under noisy channel transmissions and imperfect quantum operations. We express the capacity region as a function of the channel and network parameters (link and entanglement swap success probability), entanglement purification yield and application level parameters (target fidelity threshold). In particular, we provide necessary conditions to verify if a set of request rates belong to the capacity region of the switch. We use these conditions to find the maximum achievable end-to-end user entanglement generation throughput by solving a set of linear optimization problems. We develop a max-weight scheduling policy and prove that the policy stabilizes the switch for all feasible request arrival rates. As we develop scheduling policies, we also generate new results for computing the conditional yield distribution of different classes of purification protocols. The conclusions obtained in this work can yield useful guidelines for subsequent quantum switch designs.**

***Index Terms*—Quantum Switch; Capacity Region; Entanglement Purification; Max-weight Scheduling**

I. Introduction

Entanglement distribution is critical for many promising quantum applications such as quantum key distribution (QKD) [18], teleportation [15], quantum sensing [10], clock synchronisation [13], and distributed quantum computing [2]. Recent works [3], [4], [17], [23] and proposals [5], [14], have given much attention to distribute entanglements in a large scale quantum network that supports these applications. In an entanglement distribution network, a network of quantum switches are connected through optical fiber links. The switches are responsible for creating entanglements with their neighbors and perform a quantum operation known as *entanglement swapping* on individual entanglements to create *end-to-end user entanglements*.

In this work, we focus on a star shaped quantum network represented by a quantum switch and a set of K end user nodes as shown in Figure 1. End users are connected to the switch via a quantum communication channel. An entanglement generation source located in the middle of the channel creates maximally entangled bipartite Bell states (*Einstein-Podolsky-Rosen or EPR pairs*) and sends one half of the pair to the switch and the other to the end user. This shared entangled state between the switch and an end user is known as *link-level entanglement*. The switch and each end user can generate multiple such link-level entanglements. Requests to create end-to-end entanglement between users arrive at the switch. The role of the switch is to select specific link level entanglements to consume and perform entanglement swapping thereby generating end-to-end user entanglements to satisfy the requests. The quantum applications at the end users can then consume these end-to-end entanglements or keep them at their respective quantum memories for future usage.

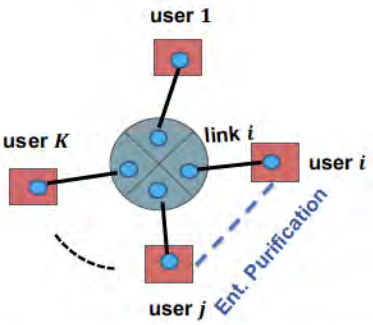


Figure 1: A quantum switch serving end-to-end entanglement to K end users.

Recently, there has been significant interest in analyzing the performance of a single quantum switch that serves multiple end users [6], [16], [20]–[22]. For example, Vardoyan et al. [20], [21] derived expressions for the maximum possible end-to-end bipartite entanglement generation rate of a quantum switch using discrete-time and continuous-time Markov chains. Nain et al. [16] derived closed form expressions for the capacity of a switch serving multi-partite entanglements to end users. They also derived the necessary and sufficient conditions for stability of the switch. However, these works [16], [20], [21] assume the creation of end-to-end entanglements in a continuous manner, i.e., once the link-level entanglements are available, they are immediately consumed to create end-to-end entanglements. The notion of serving entanglement requests from quantum applications in an on-demand manner and scheduling those requests was not considered in these works.

The request scheduling problem in a quantum switch primarily involves determining which link-level entanglements should be consumed during entanglement swapping and at

what times so that respective end user entanglement requests can be satisfied. The scheduling problem can further have multiple scheduling objectives such as maximizing the aggregate entanglement generation rate or maintaining finite request backlogs there by providing finite end-user latencies. The latter objective is also known as achieving stability for the switch and has been investigated in [6], [22] under different assumptions on lifetimes of link-level and end-to-end entanglements. In particular, Vasantam et al. [22] assumed entanglements to last for a single time slot and characterized the *capacity region* of the switch. Here, the capacity region describe the set of end-to-end request rates that the switch can stably support with finite request backlogs. Similarly, Dai et al. [6] considered infinite entanglement lifetime and proposed scheduling protocols that stabilize the switch.

While the aforementioned works on both continuous and request based models consider channel loss and entanglement swap failures, they assume the channel and quantum operations to be noiseless and do not explicitly model imperfections in capacity region calculations. Quantum transmissions are inherently noisy due to their interactions with the environment and generally produce non-maximally entangled link-level and end-to-end entanglements. Thus, the generated entanglements may not be good enough for consumption at the application level. In this work, we remove the noiseless assumption and study the problem of quantum switch scheduling with noisy channels and imperfect quantum operations. We associate a target entanglement fidelity threshold ( ) for quantum applications where fidelity is a widely used metric to quantify the quality of an entanglement. We consider the request for entanglement generation to be satisfied only when the generated end-to-end entanglement has a fidelity greater than

Due to noise in quantum channels and devices, fidelity of an entangled state decreases with each quantum operation. Let and denote the fidelity of the link-level entanglement and the entangled state created after a swap respectively. Typically, we have , where the fidelity of a perfect entangled state is considered to be . When , the generated user entanglements are not good enough for consumption at the application level. To circumvent this issue, one can perform nested entanglement purification [9] on these low quality entanglements to generate fewer number of high quality entanglements that achieve a target fidelity.

The order in which entanglement purification and entanglement swap operations are performed can lead to the following two architectures. (i) Purify Swap (PS): One where purification is performed on link-level entanglements, followed by entanglement swaps. (ii) Swap Purify (SP): One that performs entanglement swapping first on link-level entanglements and then performs purification on end-to-end entanglements. It is important to understand the trade-offs and design constraints of these architectures, which can shed some light on the future design of quantum switches. To this end, we characterize and compare the capacity regions of a quantum switch under both of these architectures.

A. Contributions

We characterize the capacity region of a quantum entanglement distribution switch under channel noise and imperfect quantum operations. Our contributions are summarized below.

- We determine the capacity region of a quantum switch as a function of the network (link swap entanglement success probability,), entanglement purification (conditional yield) and application level ( ) parameters under PS and SP architectures.
- We develop max-weight scheduling policies for both PS and SP architectures that stabilize the switch for all feasible request arrival rates.
- We generate new results for computing the conditional yield distribution of different classes of purification protocols.
- We evaluate and compare the capacity regions of PS and SP architectures under different classes of purification protocols, channel and network parameters.

The rest of the paper is organized as follows. In Section II, we briefly introduce the quantum background and the system model relevant to this work. Section III discusses some of the main results related to determining the capacity region of a switch under PS and SP architecture. Section IV presents results related to deriving the conditional yield distribution and expected conditional yield of different classes of entanglement purification protocols. We evaluate the performance of purification protocols and PS, SP architectures in Section V. Section VI concludes the paper and discusses some possible future work.

II. Technical Preliminaries

We consider a quantum switch that creates entanglements between end users as shown in Figure 1. User is connected to the switch via a fiber optic link. For brevity, we will use the term “node" to refer to both end users and the quantum switch with node referring to the switch.

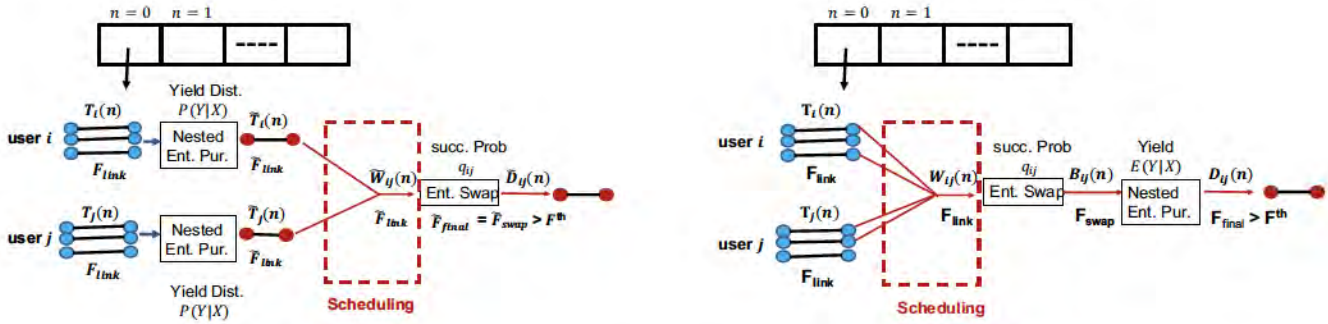
A. Quantum Background

We now briefly revisit some of the quantum operations relevant to this work.

Quantum Entanglement: A two qubit state is said to be entangled if it can not be written as a product of its individual qubit states. The following four entangled two qubit states are of particular importance.

$$\frac{}{-} \qquad \frac{}{-}$$

We refer to the two qubit entangled states and as Bell pairs and as an EPR pair shared between node and node Also, we refer to and as link level entanglement and end-to-end user entanglement respectively.



(a) (b)

Figure 2: (a) Purify and Swap architecture (b) Swap and Purify architecture in time slot n.

Fidelity: Fidelity captures the closeness between two quantum states. EPR pairs obtained after entanglement may not be perfect due to noise in quantum communication channels and quantum operations. Suppose, nodes i and j share an imperfect EPR pair $|\sigma_{ij}\rangle$. The density operator ρ_{ij} associated with $|\sigma_{ij}\rangle$ can be expressed in the Bell basis as follows.

$$\rho_{ij} = F_1 |\phi_{ij}^+\rangle \langle\phi_{ij}^+| + F_2 |\psi_{ij}^-\rangle \langle\psi_{ij}^-| + F_3 |\psi_{ij}^+\rangle \langle\psi_{ij}^+| + F_4 |\phi_{ij}^-\rangle \langle\phi_{ij}^-|, \quad (1)$$

with $\sum_{i=1}^{4} F_i = 1$. Since, we are interested in state $|\phi_{ij}^+\rangle$, F_1 is the fidelity of state $|\sigma_{ij}\rangle$. Ideally one would want to achieve a unit fidelity, i.e. value of F_1 to be equal or very close to 1. But, due to non-ideal conditions, $F_1 < 1$.

Entanglement Swapping [12]: Suppose end users i, j share EPR pairs $|\phi_{0i}^+\rangle, |\phi_{0j}^+\rangle$ with the switch. Then an entangled state $|\phi_{ij}^+\rangle$ can be created between users i and j by taking a Bell state measurement (BSM) on one qubit of each pair stored at the switch. This operation is known as entanglement swapping. Let F_{link} denote the fidelity of both link level entanglements on which entanglement swapping is being performed. Denote F_{swap} as the fidelity of the new EPR pair created by the swap. Then, F_{swap} can be expressed as an increasing function $\eta(\cdot)$ of F_{link} with $F_{swap} < F_{link}$. For example, when link level entanglements are assumed to be Werner states, $\eta(F_{link}) = 1/4 + 3/4((4F_{link} - 1)/3)^2$ [1].

Entanglement Purification [9]: The EPR pairs obtained after entanglement swapping may not be perfect and may have low fidelities. Assume that end user i and j share a number $\kappa(\geq 2)$ of imperfect EPR pairs $|\sigma_{ij}\rangle$. However, nodes i and j can perform local gates, measurements on κ $|\sigma_{ij}\rangle$ EPR pairs along with classical information exchange to obtain a single higher quality EPR pair. This operation is known as entanglement purification. Note that, entanglement purification is probabilistic. We refer interested readers to [8] for a more elaborate explanation of the purification procedure. This procedure can be repeated in a nested manner on the newly purified set of EPR pairs until desired fidelity is achieved.

K	Number of end users
F^{th}	Application level target fidelity threshold
F_{link}	Initial link level fidelity
F_{swap} ($\tilde{F}_{swap}$)	Fidelity after a successful ent. swap under SP (PS)
F_{final} ($\tilde{F}_{final}$)	Fidelity of EPR pair delivered to app under SP (PS)
p_i	Ent. generation success prob. for link i
q_{ij}	Ent. swap success prob. between user i and j
α_{max}	Max. no. of link level ent. generation attempts/slot
n	A time slot
$A_{ij}(n)$	No. of requests arrived in slot n for creating ent. b/w user i and j with Fidelity F^{th}
λ_{ij}	Avg. request arrival rate to create ent. between user i and j
$T_i(n)$	No. of successful link level ent. created for link i in slot n
$\tilde{T}_i(n)$	No. of successful purified link level ent. created for link i under PS in slot n
$W_{ij}(n)$ ($\tilde{W}_{ij}(n)$)	No. of link level ent. chosen to perform swap b/w user i and j under SP (PS) in slot n
$Q_{ij}(n)$ ($\tilde{Q}_{ij}(n)$)	No. of pending requests in slot n for creating ent. b/w user i and j with Fidelity F^{th} under SP (PS)
$B_{ij}(n)$	No. of successful ent. swaps performed in slot n b/w user i and j under SP
$D_{ij}(n)$ ($\tilde{D}_{ij}(n)$)	No. of successful ent. created in slot n b/w user i and j with fidelity $> F^{th}$ under SP (PS)
$\mathbb{E}[Y\|X]$	Conditional expected yield of an ent. pur. routine
$P_{Y\|X}$	Conditional yield distribution of an ent. pur. routine
$\eta(\cdot)$	Fidelity after ent. swap as a function of link fidelity

Table I: Summary of Notations.

B. System Model

In this Section, we introduce the system model used in the rest of the paper. We summarize notations in Table I. We use the notation $P_X(x)$ to represent the probability that the random variable X takes the value x, i.e., $P_X(x) = \Pr[X = x]$. Also, let $\mathbb{E}[X]$ denote the expected value of random variable X. We use a bold letter symbol, such as $\boldsymbol{V}$, to represent a vector.

End-to-end User Entanglement Requests: We assume time is slotted with slot duration Δ seconds. Let n denote a time slot with $n \in \{0, 1, \cdots\}$. At each time slot, requests to create $|\phi_{ij}^+\rangle$ (with $i \neq j$) arrive at the switch. Let $A_{ij}(n)$ denote the number of such requests arriving in time slot n. We assume $\{A_{ij}(n)\}_{n\geq 0}$ to be mutually independent i.i.d processes with rate λ_{ij}, i.e. $\mathbb{E}[A_{ij}(n)] = \lambda_{ij}$, $i, j \in \{1, K\}$. We also assume

that these requests require an application level minimum target fidelity[1] of .

Link Level Entanglement Generation: At the beginning of each time slot (say), the network infrastructure attempts to create number of link level entanglements between an end user and the switch. We assume the switch and the user has enough quantum memory to store all entanglements. We denote to be the success probability to create a single link level entanglement between user and the switch in each of these attempts. , where is the length of link and is its attenuation coefficient. Denote Clearly, the number of link level entanglements generated between user and the switch ( ) is a binomial random variable with parameters and , i.e.,

Let $\mathcal{A}$ denote the set of all feasible vectors of , i.e., $\mathcal{A}$. We assume the link level entanglements to be valid only for a single time slot and they decohere at the end of the time slot. Denote to be the fidelity of a successfully generated link level entanglement. Due to noise induced in the link, .

Nested Entanglement Purification: When the fidelities of the end-to-end entanglements generated after swaps are less than , the entanglements become unusable for consumption at the application level. To circumvent this issue, one can perform nested entanglement purification on these low quality entanglements to generate entanglements with a target fidelity. Suppose denotes the random variable representing the total number of output EPR pairs with fidelity greater than produced by a nested purification routine by consuming low quality EPR pairs. We refer to and as the conditional yield[2] distribution and the conditional expected yield of the nested entanglement purification routine. We derive expressions for and for different classes of nested entanglement purification protocols in Section IV. Such a purification can be performed before or after entanglement swapping routine resulting in the following two architectures as shown in Figure 2 (a) and (b).

(a) **Purify and swap (PS) architecture -** In this architecture, nested entanglement purification is first performed on individual link level entanglements after which entanglement swapping is performed.
(b) **Swap and purify (SP) architecture -** This architecture first performs entanglement swapping on link level entanglements followed by end-to-end purification.

[1]Throughout the paper, the fidelity of a state is always with respect to the ideal EPR pair .

[2]Note that, the definition of yield is slightly different from what has been used in literature. Previous work refers to as the yield of a purification routine.

Let ( ) denote the number of high quality end-to-end entanglements that are delivered to the requesting application at time slot each with fidelity ( ) such that ( ) in the SP (PS) architecture.

Entanglement Scheduling: At the beginning of each time slot, the switch stores the unserved requests in a queue. Let ( ) denote the queue size in time slot in the SP (PS) architecture. Also, let and . Similarly, let . The switch makes scheduling decisions in each time slot and selects ( ) number of link level entanglements to perform entanglement swapping under SP (PS). Since entanglement swapping procedures are probabilistic (with success probability), only a fraction of ( ) EPR pairs are created between user and at time slot under SP (PS) architecture. Denote , and .

III. Capacity Region Computation: Main Results

In this section, we present the main results related to the capacity region computation of a switch under both PS and SP architecture. We also propose corresponding max weight scheduling policies that stabilize the switch for all feasible request arrival rates. We relegate the proofs in this section to the Appendix VIII.

Definition 1: A switch is defined to be stable under a scheduling policy if the process converges in distribution to independent of the initial condition and [19].

Definition 2: The capacity region of a switch is defined to be the set of request rates for which the switch remains stable.

A. PS Architecture

In this architecture, entanglement purification is first performed on link level entanglements, followed by entanglement swap. We denote and to be the number of link-level entanglements and their fidelities after link-level purification is performed. Since, the purification is on the link level, we need to identify a target threshold for fidelity on link-level ( ). We set . We define the probability distribution of as

(2)

where is the conditional yield distribution of the purification protocol. The following theorem computes the capacity region under PS architecture.

Theorem 1: The capacity region of the switch under PS is given by,

s.t.

$\mathcal{A}$

$\mathcal{A}$ (3)

where is defined as the component-wise inequality between vectors and

Intuitively, can be interpreted as the fraction of time in which a scheduling policy selects given . Similarly, is the element-wise multiplication of the vectors and . can be thought of as the average number of successfully served entanglement requests between users and in time slot . Note that, the noiseless capacity region, i.e., the case when no purification is applied and there is no notion of can be obtained by setting . This is precisely what authors in [22] have obtained for the case when
We now define the following max-weight scheduling policy for the PS architecture.

Definition 3: (MW-PS) Under this policy, the switch selects the purified link level entanglements from the available entanglements ( ) in the following manner to perform entanglement swaps at each time slot .

s.t (4)

The theorem below discusses the stability properties of MW-PS.

Theorem 2: If is finite , then the capacity region of MW-PS coincides with

B. SP Architecture

In this architecture, entanglement swapping is performed first, followed by end-to-end purification. We characterize its capacity region as follows.

Theorem 3: The capacity region of a quantum switch under SP is given by

s.t.

$\mathcal{A}$

$\mathcal{A}$ (5)

Also, is given by the following expression.

Here, is the conditional expected yield value of a purification protocol.

The quantity can be interpreted as the average number of successfully served requests in time slot after purification. We now define a max weight scheduling policy corresponding to SP.

Definition 4: (MW-SP) In this policy, the switch selects the link level entanglements from the set of available link entanglements ( ) in the following manner to perform entanglement swap at each time slot .

s.t (6)

Theorem 4: If is finite , then the capacity region of MW-SP coincides with

C. Boundary of the Capacity Region and Maximum Throughput

We now formulate a weighted throughput maximization problem to determine the boundary of and as follows. Given a quantum switch, we want to find the maximum weighted achievable user request rate that stabilizes it. Let denote the weight associated with end user-pair . Formally, we define the following optimization problem for PS:

s.t

(7)

A similar optimization problem can be formulated for SP architecture as well. One application of the above optimization problem is to sketch out the capacity region of the switch by solving it for different instances of , as each solution falls on the boundary of the capacity region. Note that, both the objective and constraints in (7) are linear and involves continuous decision variables. Hence optimization problem (7) corresponds to a linear program.

IV. Properties of Specific Nested Entanglement Purification Protocols

We now derive the conditional yield distribution and conditional expected yield of the following classes of nested entanglement purification protocols. Let denote the random variable representing the total number of input EPR pairs submitted for purification to a nested purification routine. Similarly, let denote the random variable representing the the total number of output EPR pairs with fidelity greater than produced by the nested purification routine. We are interested in and .

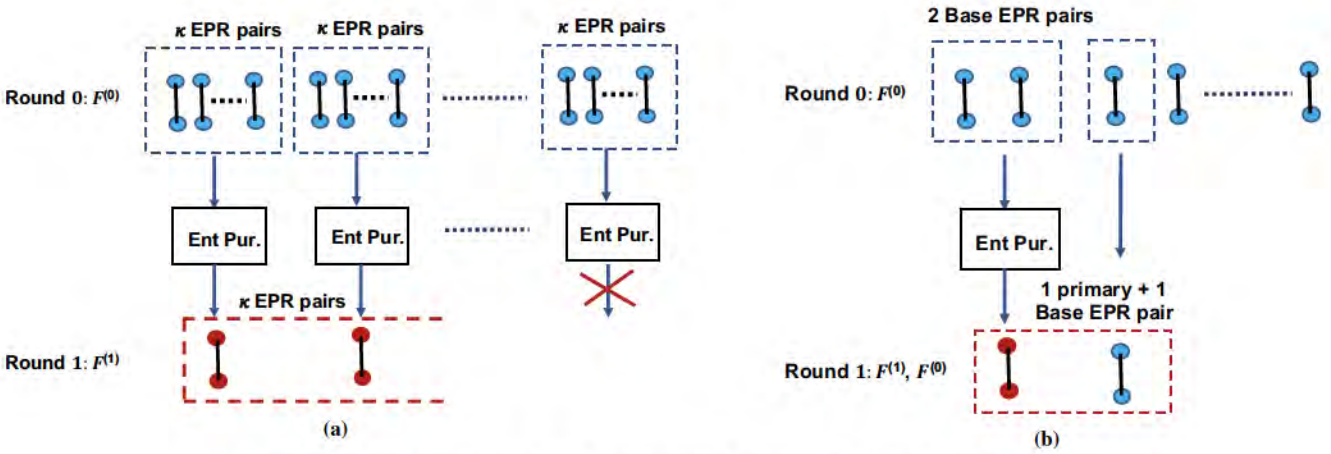


Figure 3: (a) Symmetric and (b) Entanglement Pumping Purification Protocols.

Protocol	Input State	κ	$\omega_S(F)$	$\chi_S(F)$
DEJMPS [7]	Bell-Diagonal (See (1) with $F_1 = F$)	2	$(F+F_2)^2 + (F_3+F_4)^2$	$\frac{F^2+F_2^2}{(F+F_2)^2+(F+F_4)^2}$
DEJMPS [7]	Binary	2	$F^2 + (1-F)^2$	$\frac{F^2}{F^2+(1-F)^2}$
BBPSSW [1]	Werner	2	$\left(F+\frac{1-F}{3}\right)^2 + \left(\frac{2(1-F)}{3}\right)^2$	$\frac{F^2+\frac{1}{9}(1-F)^2}{F^2+\frac{2}{3}F(1-F)+\frac{5}{9}(1-F)^2}$

Table II: Success probability and output fidelity for different symmetric Purification protocols.

A. Symmetric Purification Protocols

In symmetric purification protocols, purification is performed on identical EPR pairs, i.e., all input EPR pairs have equal fidelities. The protocol is often performed in multiple rounds, hence also known as recurrence protocol [9]. At each purification round, the total number of available EPR pairs are divided into groups of $\kappa > 1$ number of EPR pairs as shown in Figure 3 (a). Purification is then performed on each group. If the purification is successful, then only one EPR pair out of κ EPR pairs are kept in each group. If the purification is unsuccessful, all of the κ EPR pairs are discarded. The remaining EPR pairs are used for purification at the next round.

At each round, fidelities of EPR pairs increase by a certain fraction after purification. Purification is continued until either there are no EPR pairs left or the application level fidelity threshold is reached for the output EPR pairs. We denote L as the maximum number of purification rounds.

Let $Y^{(l)}$ and $F^{(l)}$ denote the total number of output EPR pairs and their fidelities after l rounds of purification for $l = 1, 2, \cdots, L$. Thus we have $F^{(L)} \geq F^{th}$ but $F^{(L-1)} < F^{th}$. We also have $Y^{(L)} = Y$ and $Y^{(0)} = X$.

Let x denote the total number of input EPR pairs, each with fidelity F, available at the beginning for purification. At round l, each purification unit takes κ number of EPR pairs each with fidelity $F^{(l-1)}$ and produce one EPR pair of fidelity $F^{(l)} = \chi_S(F^{(l-1)})$ with $F^{(l-1)} < F^{(l)}$ and $l \in \{1, 2, \cdots, L\}$, $F^{(0)} = F$. Here, $\chi_S(\cdot)$ is the output fidelity function.

Purification is probabilistic and its success probability depends on the fidelity of the input EPR pairs. Let $r_S^{(l)}$ denote the success probability of a purification routine in round l. Thus we have

$$r_S^{(l)} = \omega_S(F^{(l-1)}) \quad \forall l \in \{1, 2, \cdots, L\}, \tag{8}$$

where $\omega_S(\cdot)$ denotes the success probability as a function of fidelity of input EPR pairs.

The probability distribution of number of output EPR pairs produced after l rounds of purification can be derived by the following recursion with $\forall l \in \{2, \cdots, L\}$.

$$P_{Y^{(l)}|X}(y|x) \tag{9}$$

$$= \sum_{z=y*\kappa}^{\lfloor x/\kappa^{l-1} \rfloor} \binom{\lfloor z/\kappa \rfloor}{y} \left(r_S^{(l)}\right)^y \left(1 - r_S^{(l)}\right)^{\lfloor z/\kappa \rfloor - y} \cdot P_{Y^{(l-1)}|X}(z|x), \tag{10}$$

with $P_{Y^{(1)}|X}(y|x) = \binom{\lfloor x/\kappa \rfloor}{y} (r_S^{(1)})^y (1 - r_S^{(1)})^{\lfloor x/\kappa \rfloor - y}$. The conditional yield distribution is given by $P_{Y|X}(y|x) = P_{Y^{(L)}|X}(y|x)$. The conditional expected yield function is given by $\mathbb{E}[Y|X = x] = \sum_{y=1}^{\lfloor x/\kappa^L \rfloor} y * P_{Y^{(L)}|X}(y|x)$.

We present the success probability and output fidelity functions of different protocols that perform $\kappa : 1$ purification in Table II. Note that, depending on the input state of an EPR pair, $\omega_S(\cdot)$ and $\chi_S(\cdot)$ may be different. Consider the density operator representation of an input EPR pair as shown in Equation (1). When $F_1 = F$ and $F_2 = F_3 = F_4 = (1-F)/3$, we call the entangled state a Werner state (White noise). When $F_1 = F$ and only one of F_2, F_3 or F_3 is non-zero, i.e., equals to $(1 - F)$, the state is called a Binary state. For example: if the channel has bit flip errors, then a particular kind of binary state ($F_1 = F, F_2 = F_4 = 0, F_3 = 1 - F$) is generated.

B. Entanglement Pumping

Entanglement Pumping [9] is an asymmetric purification protocol where one EPR pair is purified using multiple auxiliary EPR pairs until a desired target fidelity is achieved. Then the process repeats itself to produce more high quality EPR pairs until all auxiliary EPR pairs are exhausted. The protocol is shown in Figure 3 (b). We call the EPR pair that goes under purification as *primary* and other auxiliary pairs as *base* EPR pairs.

We start with a given number of base EPR pairs with Fidelity We select the first base EPR pair to be a primary EPR pair and purify it using the second base EPR pair. If the purification is successful, the fidelity of the primary EPR pair increases by a fraction and the base EPR pair gets sacrificed. The primary EPR pair then repeatedly gets purified by sacrificing other base EPR pairs until its fidelity is above . If a purification fails at any step, the next primary EPR pair is chosen from the set of remaining base EPR pairs and the purification process repeats from the beginning.

Similar to the definitions in Section IV-A, let denote the fidelity of a primary EPR pair after rounds of purification. Suppose rounds of purification is needed for a primary EPR pair to achieve the target fidelity. Since entanglement pumping is asymmetric, the fidelities of input EPR pairs for a single purification operation are different. We denote and to be the output fidelity of a primary EPR pair and purification success probability functions respectively. Thus we have, We denote to be the purification success probability for round i.e.

We use the theory of discrete-time renewal processes [11] to derive the distribution for entanglement pumping protocol as follows. Henceforth, we will call the purified primary EPR pair with fidelity greater than as a high quality EPR pair. We refer to the event of generating a high quality EPR pair as a renewal event. We sequentially select a base pair to purify the primary EPR pair and the ids of base EPR pairs can be mapped to (discrete) time in a renewal process. The number of EPR pairs sacrificed to generate one high quality EPR pair can be considered as the renewal inter-occurrence time. Let denote such renewal inter-occurrence times with distribution . Thus, by the theory of renewal processes [11], the conditional expected yield is given by the following recursion.

(11)

with .

Now, let denote renewal event occurrence times, i.e. can be found out by fold convolution using the recursion, The conditional yield distribution can be computed as

(12)

Finally, the distribution can be computed by the following recursion.

(13)

Here, . The set in the indicator function denotes the infeasibility set of a purification scheme. For example, for a purification scheme with , i.e. the set of odd numbers.

V. Numerical Results

In this section we perform a numerical study to compare the performance of different architectures and purification protocols. In order to obtain the boundary of the capacity region, we solve optimization problem (7) using the IBM CPLEX solver. The goals of this study are to (i) compare the performance of PS and SP architectures, (ii) to determine which purification protocol works best with respect to switch stability, and (iii) the effect of different network and application-level parameters on the capacity region of the switch. We first present the experimental setup and then discuss the results.

Parameter		max	i	ij	$link$	th
Value	3	10	0 9	0 9	0 90	0 85

Table III: Parameters used for numerical studies.

We consider three end users connected to a quantum entanglement distribution switch with request arrival vector and We assume the distance between the switch and an end user to be km with fiber attenuation coefficient 0.2 dB/km [6]. Unless specified, we use the parameters presented in Table III for numerical simulations.

Comparison of PS and SP architectures: We compare the capacity region of PS and SP under the DEJMPS protocol as shown in Figure 4 (a). We observe that PS has a larger capacity region compared to SP. This advocates that the switch under PS is stable for a wider range of end-user entanglement generation demands and thus, more robust to unexpected demand fluctuations. We obtain similar results for other purification protocols listed in Table II and hence, we omit them. Hence forth, for all remaining experiments, we will use the PS architecture to determine the capacity regions. We also plot the noiseless capacity region, which serves as an upper bound on the maximum capacity that can be achieved by any purification scheme under any architecture. The noiseless capacity region is obtained for the setting where there is no notion of and links as well as quantum operations are assumed to be perfect. Note that, the noiseless capacity is significantly higher than those achieved by any of the architectures. The gap between noiseless capacity and

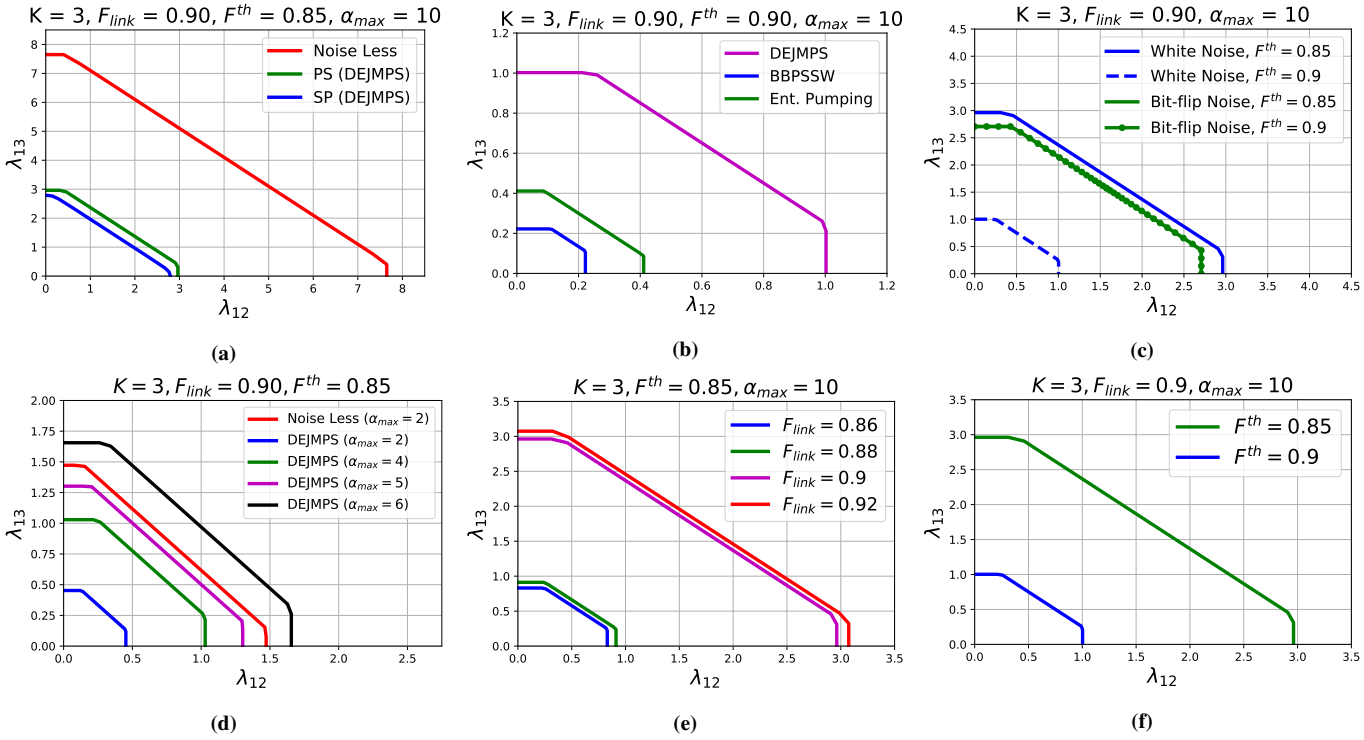


Figure 4: Capacity Regions computed for different architectures, network parameters and purification protocols. (c)-(f) are computed for PS architecture and DEJMPS purification protocol.

capacities of purification architectures in Figure 4(a) suggests that there could be room for improvement in purification protocol performance.

Shape of the Capacity Region: We now explicate the reason behind the common shape of the capacity regions as shown in Figure 4. We explain the shape of the noiseless capacity region in 4 (a), while the same arguments hold for other plots. Assume that requests of type (type) are satisfied using links and (links and). Let denote the maximum supportable rate between end-users and when . Now when , but sufficiently small, these requests of type can still be served using link-level entanglements of links and whenever link-level entanglements of link are not available. When link-level entanglements of link are available, the switch can only serve requests of type and stably support a rate . When both and are sufficiently large, there lies a competition for link-level entanglements of link to serve either type or type requests. Hence, in this case, we obtain the diagonal line.

Comparison of Purification Protocols: We now compare the capacity regions achieved by three purification protocols: two symmetric (DEJMPS, BBPSSW) and one asymmetric (Pumping). In entanglement pumping, the functions and are computed using an asymmetric version of DEJMPS protocol [9]. We set . In Figure 4 (b), we observe that DEJMPS protocol outperforms the other protocols. DEJMPS requires fewer purification rounds ( ) than BBPSSW to reach a target fidelity. As the number of base EPR pairs that are sacrificed grows exponentially with , DEJMPS achieves a larger capacity region.

Bit-flip Noise versus White Noise: We compare the performance of DEJMPS protocol under two different noise models: Bit-flip noise and White noise as shown in Figure 4 (c). Short descriptions of both models is found in Section IV-A. We consider two settings: and . When , both noise models require one ( ) purification round. In this setting, DEJMPS under White noise produces lower quality EPR pairs with higher success probability compared to Bit-flip model. Since, we are only interested in the yield, the success probability component dominates in capacity region calculations. Thus DEJMPS under White Noise provides a slightly larger capacity region as compared to that of under Bit-flip noise. However, for a more stringent requirement on ( ), and for Bit-flip and White noise models respectively. The number of EPR pairs that are sacrificed grows exponentially with , resulting in a significantly smaller capacity region under White noise model. We conclude that White noise is difficult to deal with under stringent application level quality requirements.

Effect of : In Figure 4(d), we study the effect of (number of link-level entanglement attempts) on the capacity region of the switch. We compute the noiseless capacity for and compare it to the noisy capacity for different values of starting with We observe that the noiseless capacity is significantly larger than the noisy capacity for the same value of . As expected, increasing increases the noisy capacity. We also observe that one

has to make at least three times more link-level entanglement attempts (black curve) to achieve the noiseless capacity.

Effect of and : We study the effect of and on the capacity region of the switch in Figures 4 (e) and (f) respectively. With an increase in the quality of initial link-level entanglements, larger capacity regions are obtained as shown in Figure 4 (e). Note that, the improvement in performance is minute between and but significant between and . This can be explained by the fact that when and when Thus the transition from to involves one less number of purification round resulting in significant gain in terms of capacity. We also compare the capacity regions of the switch for different values of in Figure 4 (f). We observe that the capacity region reduces drastically with stringent application level quality requirements.

VI. Conclusion and Open Problems

In this work, we analyzed the effect of channel noise and entanglement purification on the capacity region of a quantum switch. Going further, we would like to extend our analysis to an arbitrary network setting. Conceptually, this seems doable provided that there is a predefined path in place for each user pair. However, the number of optimization variables in (7) can grow rapidly as a function of numbers of nodes and user pairs. Also, the focus of this work has been on a swap-purify architecture and a purify-swap architecture. It would be interesting to consider an alternate purify-swap-purify architecture where link-level entanglements are purified followed by swaps followed by end-to-end purification. For the network setting, one can of course perform multi-hop purification, then do swapping followed by again multi-hop purification.

VII. Acknowledgments

This research was supported by the NSF ERC for Quantum Networks (CQN), awarded under cooperative agreement number 1941583 and by NSF grant CNS-1955834.

VIII. Appendix

A. Proof of Theorem 1 and Theorem 2

The proofs for the PS architecture are very similar to that obtained in [22] for . When and there is no notion of noise, the random variable is Bernoulli distributed. When and purification is applied, follows the distribution as mentioned in Equation (2). Following the capacity calculation and stability proof of max-weight protocol in [22] with the distribution of as expressed in (2) yields Theorem 1 and 2. We now outline the proofs for the SP architecture.

B. Proof of Theorem 3

We assume that link level entanglements decohere after one time slot. Hence, only number of entanglement swaps are performed in time slot . Thus, . Also, where and Let where the function captures the effect of purification. evolves as . We start with the assumption that the process is an irreducible Markov chain. The conditions that a scheduling policy should satisfy for this assumption can be found in [19], [22]. For the process to be stable, the following condition should hold true.

$$\mathcal{A} \tag{14}$$

We now derive for all as follows.

$$\tag{15}$$

and
Therefore,

$$\tag{16}$$

Substituting (16) in Equation (14) and removing the conditioning on in Equation (14) as discussed in [22], we get the capacity region derived in Theorem 3.

C. Proof of Theorem 4

To prove this, we apply Lyapunov stability of Markov chains using the Lyapunov function
We want to show that
. Thus we have

$$\tag{17}$$

It is easy to show that,
, for some constant Now we simplify,

$$\tag{18}$$

Substituting (16) in (18) and proceeding in a similar manner as in [22] establishes the Lyapunov stability condition.

REFERENCES

[1] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller. Quantum repeaters: The role of imperfect local operations in quantum communication. *Phys. Rev. Lett.*, 81:5932–5935, Dec 1998.

[2] A. Broadbent, J. Fitzsimons, and E. Kashefi. Universal blind quantum computation. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*. IEEE, oct 2009.

[3] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner. Distributed routing in a quantum internet. *arXiv preprint arXiv:1907.11630*, 2019.

[4] C. Cicconetti, M. Conti, and A. Passarella. Request scheduling in quantum networks. *IEEE Transactions on Quantum Engineering*, 2:2–17, 2021.

[5] A. Dahlberg, M. Skrzypczyk, T. Coopmans, L. Wubben, F. Rozpędek, M. Pompili, A. Stolk, P. Pawełczak, R. Knegjens, J. de Oliveira Filho, et al. A link layer protocol for quantum networks. In *Proceedings of the ACM Special Interest Group on Data Communication*, pages 159–173. 2019.

[6] W. Dai, A. Rinaldi, and D. Towsley. Entanglement swapping in quantum switches: Protocol design and stability analysis, 2021.

[7] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Phys. Rev. Lett.*, 77:2818–2821, Sep 1996.

[8] W. Dür and H. J. Briegel. Entanglement purification and quantum error correction. *Reports on Progress in Physics*, 70(8):1381–1424, jul 2007.

[9] W. Dür, H. J. Briegel, J. I. Cirac, and P. Zoller. Quantum repeaters based on entanglement purification. *Physical Review A - Atomic, Molecular, and Optical Physics*, 59(1):169–181, 1999.

[10] Z. Eldredge, M. Foss-Feig, J. A. Gross, S. L. Rolston, and A. V. Gorshkov. Optimal and secure measurement protocols for quantum sensor networks. *Phys. Rev. A*, 97:042337, Apr 2018.

[11] W. Feller. An introduction to probability theory and its applications.

[12] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. "event-ready-detectors" bell experiment via entanglement swapping. *Phys. Rev. Lett.*, 71:4287–4290, Dec 1993.

[13] P. Komar, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin. A quantum network of clocks. *Nature Physics*, 10(8):582–587, 2014.

[14] S. Lloyd, J. H. Shapiro, F. N. Wong, P. Kumar, S. M. Shahriar, and H. P. Yuen. Infrastructure for the quantum internet. *ACM SIGCOMM Computer Communication Review*, 34(5):9–20, 2004.

[15] X.-S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin, and A. Zeilinger. Quantum teleportation over 143 kilometres using active feed-forward. *Nature*, 489(7415):269–273, 2012.

[16] P. Nain, G. Vardoyan, S. Guha, and D. Towsley. On the analysis of a multipartite entanglement distribution switch. *Proc. ACM Meas. Anal. Comput. Syst.*, 4(2), jun 2020.

[17] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha. Routing entanglement in the quantum internet. *npj Quantum Information*, 5(1):1–9, 2019.

[18] P. W. Shor and J. Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, Jul 2000.

[19] L. Tassiulas. Scheduling and performance limits of networks with constantly changing topology. *IEEE Transactions on Information Theory*, 43(3):1067–1073, 1997.

[20] G. Vardoyan, S. Guha, P. Nain, and D. Towsley. On the exact analysis of an idealized quantum switch. *Performance Evaluation*, 144:102141, 2020.

[21] G. Vardoyan, S. Guha, P. Nain, and D. Towsley. On the stochastic analysis of a quantum entanglement distribution switch. *IEEE Transactions on Quantum Engineering*, 2:1–16, 2021.

[22] T. Vasantam and D. Towsley. A throughput optimal scheduling policy for a quantum switch. In *Quantum Computing, Communication, and Simulation II*. SPIE, mar 2022.

[23] Y. Zhao, G. Zhao, and C. Qiao. E2e fidelity aware routing and purification for throughput maximization in quantum networks. In *Proceedings of the IEEE INFOCOM*, 2022.